

Fig. 1
1994

RESULT 5

PCPAL3 2403 bp mRNA linear PLN 07-DEC-1994
LOCUS P.crispum mRNA for phenylalanine ammonia-lyase 3.
DEFINITION X81159
ACCESSION X81159.1 GI:535007
KEYWORDS deaminase; PAL-3 gene; phenylalanine ammonia-lyase.
SOURCE Petroselinum crispum.
ORGANISM Petroselinum crispum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids II; Apiales; Apiaceae; Petroselinum.
1 (bases 1 to 2403)
Appert, C., Logemann, E., Hahlbrock, K., Schmid, J. and Amrhein, N.
Structural and catalytic properties of the four phenylalanine
ammonia-lyase isoenzymes from parsley (Petroselinum crispum Nym.)
Eur. J. Biochem. 225 (1), 491-499 (1994)

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
95010141
7925471
2 (bases 1 to 2403)
Appert, C.
Direct Submission
Submitted (22-AUG-1994) C. Appert, Institut fuer
Pflanzenwissenschaften, Eidgenossische Technische Hochschule,
Zuerich, Universitaetstrasse 2, 8092 Zurich, SWITZERLAND
Location/Qualifiers

FEATURES
source
1..2403
/organism="Petroselinum crispum"
/db_xref="taxon:4043"
89..2245
/gene="PAL3"
89..2245
/gene="PAL3"
/EC_number="4.3.1.5"
/note="tetramere subunit"
/product="phenylalanine ammonia-lyase 3"
/protein_id="CA57057.1"
/db_xref="GI:535008"
/db_xref="SWISS-PROT:P45729"
/translation="MAYVNGITNGHANGNGLDGCKKEDPLNMGVAAEAATGSHLDEV
KRMVAEYKRPVVKLEGETLITISOVAISARDGSGVVELSEARAGVAKASDWVDSM
NKGDSYGVTTGFGATSHRTKGGALQKELIRFLNAGIFGSGAEGNNPLPSATRA
ANLVIRINTLOGYSIGIREILEITKELFNHNTPCPLRGITIASGDLVPLSYIAGLL
TGRPSKAVGPTGVTISPSEAFKLAGVGGFEELQPKREGIALVNGTAVGSGMASVYLE
ENNTLAVLAEVMSAIFAEVMGCKRPETDHLTKLKHHPGOIEMAAIMHEHLDGSAYVK
AAOKLHMDPLQKPKODRYALRTSPOMLGPOIETIRSTKMIEREINSVNDPLIDVS
RNKAIHGNGFQSPIGVSMNTRFLAIAITGKLMAPQSELVNDPTNGDPSNLSGGN
PSIDYGFKAELIAVMSYCELOFLANPVTNHOVSAEQHNDVNSLGLISSRTSEAVE
IKIMSTPLVGLCOAIDLRHLEENLKSTVKNVSAVAKRYLTWVGNGELHPSRPECVE
DLRYVDREYITPAVIDPCSATYPLMOKLRETLVEHALLNNGDKERNLSTSIQKIAAF
EDBLKALLPKREYETARALESQNPAPIRIKECRSYPLKRVREELCTEYLTGKVAS
PCBEFEKVTAMSGEITIDPLECLESNGAPLPIC"

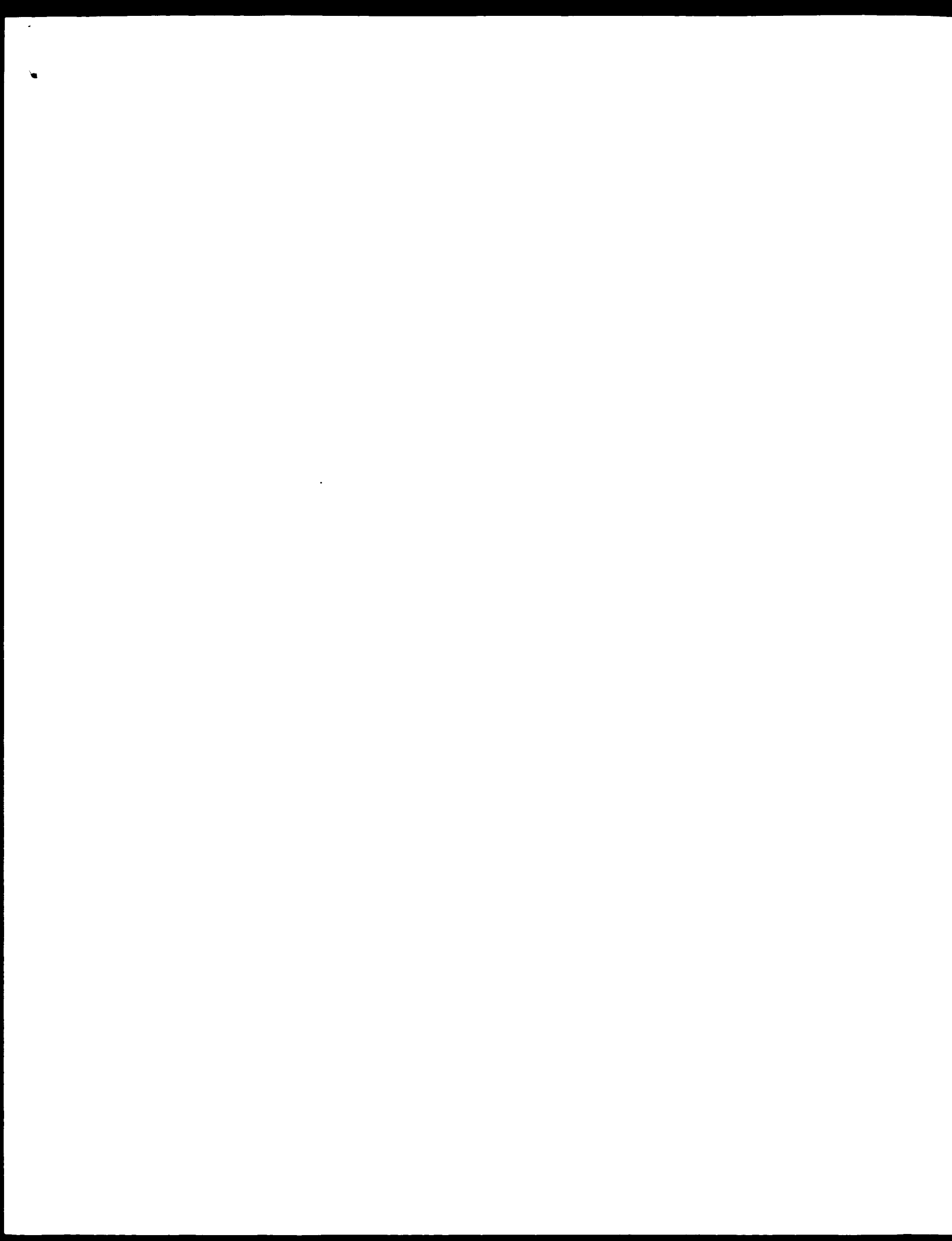
BASE COUNT 708 a 461 c 573 g 661 t
ORIGIN

Query Match 52.0%; Score 1270.8; DB 8; Length 2403;
Best Local Similarity 74.2%; Pred. No. 11e-293;
Matches 1634; Conservative 0; Mismatches 562; Indels 6; Gaps 2;



135 ACCTAATGAGCTGCTTAATGAGTGTGCATCAAGCATCCATTGAACCTGGGAGTTGCAG 194
126 ACAGGAAAGGATTGATTTGTGCTATGAGAGAGAAATCCTTTGAACCTGGGAGTGGCTG 185
195 CGAGAGCGTTGACCGGAAAGTCACTTGTATGAGTGAAGAAAGTGGTGGGCTTCAGAA 254
186 CGAGAGCATTTGACAGGAGTCACTTGTATGAGTGAAGAAAGTGGTGGGCTTCAGAA 245
255 AGCGGCTGTGAAGTCCGAG 314
246 AGCGGCTGTGAAGTCCGAG 305
315 CTGCTATGACATGACACCGTGAAGTGAAGTGGTGGGCTTCAGAAAGTGGGCTTCAGAA 374
306 CTGAGGATGATGAGTGGTGGGCTTCAGAAAGTGGGCTTCAGAAAGTGGGCTTCAGAA 362
375 AGCGGCTGTGAAGTCCGAG 434
363 AGCGGCTGTGAAGTCCGAG 422
435 CCAGCGCTTGGGCGGCTTCAG 494
423 CTGAGGATGATGAGTGGTGGGCTTCAGAAAGTGGGCTTCAGAAAGTGGGCTTCAGAA 482
495 AGCTCATGATGATTTGAAG 551
483 AGCTCATGATGATTTGAAG 542
552 CACTTCCATGATGATTTGAAG 611
543 CATTACCACTGATTTGAAG 602
612 GTTACCGCGGATGATTTGAAG 671
603 GCTATTCAGAGATGATTTGAAG 662
672 TCACCGCTTGGTGGGCTTCAG 731
663 TTAAGCTTGGTGGGCTTCAG 722
732 CATACATGCGCGGCTTCAG 791
723 CCAACATGCGCGGCTTCAG 782
792 AGCTCATGATGATTTGAAG 851
783 TAACTCATGATGATTTGAAG 842
852 TACAGCGGATGATTTGAAG 911
843 TACAGCGGATGATTTGAAG 902
912 CCATGCTTGGTGGGCTTCAG 971
903 CTATGCTTGGTGGGCTTCAG 962
972 TCAGCGGATGATTTGAAG 1031
963 TCAGCGGATGATTTGAAG 1022
1032 ATACCGCTTGGTGGGCTTCAG 1091
1023 ACATTCGCGGCTTCAG 1082
1092 ACCTCAAGCGGCTTCAG 1151
1083 ACCTCAAGCGGCTTCAG 1142
1152 GTTATGCTTGGTGGGCTTCAG 1211
1143 GATATGCTTGGTGGGCTTCAG 1202

1212 CAACGAAATGATTCAG 1271
1203 CAACGAAATGATTCAG 1262
1272 CAACGAAATGATTCAG 1331
1263 CAACGAAATGATTCAG 1322
1332 CAACGAAATGATTCAG 1391
1323 CAACGAAATGATTCAG 1382
1392 CAACGAAATGATTCAG 1451
1383 CAACGAAATGATTCAG 1442
1452 CAACGAAATGATTCAG 1511
1443 CAACGAAATGATTCAG 1502
1512 CAACGAAATGATTCAG 1571
1503 CAACGAAATGATTCAG 1562
1572 CAACGAAATGATTCAG 1631
1563 CAACGAAATGATTCAG 1622
1632 CAACGAAATGATTCAG 1691
1623 CAACGAAATGATTCAG 1682
1692 CAACGAAATGATTCAG 1751
1683 CAACGAAATGATTCAG 1742
1752 CAACGAAATGATTCAG 1811
1743 CAACGAAATGATTCAG 1802
1812 CAACGAAATGATTCAG 1871
1803 CAACGAAATGATTCAG 1862
1872 CAACGAAATGATTCAG 1931
1863 CAACGAAATGATTCAG 1922
1932 CAACGAAATGATTCAG 1991
1923 CAACGAAATGATTCAG 1982
1992 CAACGAAATGATTCAG 2051
1983 CAACGAAATGATTCAG 2042
2052 CAACGAAATGATTCAG 2111
2043 CAACGAAATGATTCAG 2102
2112 CAACGAAATGATTCAG 2171
2103 CAACGAAATGATTCAG 2162
2172 CAACGAAATGATTCAG 2231
2163 CAACGAAATGATTCAG 2222
2232 CAACGAAATGATTCAG 2291
2223 CAACGAAATGATTCAG 2282
2292 CAACGAAATGATTCAG 2351
2283 CAACGAAATGATTCAG 2342



GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 28, 2003, 19:53:15 ; Search time 6234 Seconds

(without alignments)
11400 233 Million cell updates/sec

Title: US-09-964-992a-3

Perfect score: 2442
Sequence: 1 gagcaatcgtatcaatcccc.....aaaaaaaaaaaaaaaaaaaaa 2442

Scoring table: IDENTITY_NMC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: GenBank:

1: gb-ba:
2: gb-ba:
3: gb-in:
4: gb-cm:
5: gb-cv:
6: gb-pat:
7: gb-ph:
8: gb-pl:
9: gb-pr:
10: gb-pr:
11: gb-pr:
12: gb-sy:
13: gb-un:
14: gb-vi:
15: em-bo:
16: em-fun:
17: em-hum:
18: em-in:
19: em-mu:
20: em-mu:
21: em-or:
22: em-ov:
23: em-pat:
24: em-ph:
25: em-pl:
26: em-ro:
27: em-rls:
28: em-un:
29: em-vi:
30: em-hq-hum:
31: em-hq-inv:
32: em-hq-other:
33: em-hq-mus:
34: em-hq-pla:
35: em-hq-rod:
36: em-hq-mam:
37: em-hq-vrt:
38: em-sy:
39: em-hqo-hum:
40: em-hqo-pla:
41: em-hqo-other:

Prod. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	2442	100.0	2442	8	AF299330 Lactuca s
2	1550.2	63.5	2322	8	Y11461 Helianthus
3	1286.2	52.7	2444	8	AF401636 Rehmannia
4	1280.6	52.4	2375	8	X17462 P. cristum m
5	1270.8	52.0	2403	8	PCPAIL4
6	1259.2	52.0	2448	8	AF126116 Agastache
7	1254.2	51.6	2448	8	NTPEPAL
8	1250.8	51.4	2330	8	CMEPAL
9	1247.4	51.2	2527	8	TOBPAIL
10	1247.4	51.1	2495	8	AB042520
11	1235.2	50.6	2428	8	SSNVAL
12	1234	50.6	2403	8	DLJ002221
13	1231.4	50.4	2396	8	ICPPAILA
14	1227.8	50.3	2384	8	ICPPAIL2
15	1227.4	50.3	2441	8	D83075
16	1220.4	50.0	2352	8	D83076
17	1219.6	49.9	2380	8	POPPALCA
18	1219	49.9	2414	8	AF092957
19	1217.2	49.8	2439	8	AF337955
20	1215.8	49.8	2406	8	AF367308
21	1210.2	49.6	2527	8	AF045919
22	1208.8	49.5	2154	8	AV134695
23	1206.2	49.4	2335	8	IPMVAL
24	1204.6	49.3	2309	8	AY079363
25	1203.2	49.3	2178	6	AX412245
26	1203.2	49.3	2178	6	AX412319
27	1203.2	49.3	2178	6	AX412320
28	1202.8	49.3	2409	8	MSPAIL
29	1201.6	49.2	2253	8	AY036011
30	1200.6	49.2	2402	8	AF088956
31	1199.2	49.1	2448	8	AF036948
32	1194.6	48.8	2607	8	CLU43338
33	1178	48.2	2427	8	CCL248753
34	1173.2	48.0	2471	8	CCL238754
35	1170.4	47.9	2367	8	PEAPAIL
36	1159	47.5	2607	6	E04042
37	1158.6	47.4	2276	8	AF250846
38	1150.2	47.1	2501	8	AF337954
39	1102.6	45.2	2536	8	AY096657
40	1093.8	44.8	2249	8	BPPAIL
41	1089.4	44.6	3465	8	D85850
42	1081.8	44.3	2044	8	PA016130
43	1080	44.2	R0917	8	AF004502
44	1067.6	43.7	4472	8	PEAPAIL2
45	1061.4	43.5	6306	8	AB008200

ALIGNMENTS

RESULT 1
LOCUS AF299330 2442 bp mRNA linear PLN 31-DEC-2001
DEFINITION Lactuca sativa phenylalanine ammonia-lyase mRNA, complete cds.
ACCESSION AF299330
VERSION AF299330.1 GI:16001006
KEYWORDS
SOURCE
ORGANISM
Lactuca sativa.
Lactuca sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; easterids II; Asterales; Asteraceae; Lactuceae;
Lactuca.
1 (bases 1 to 2442)
REFERENCE
AUTHORS Campos-Vargas,R., Nonogaki,H., Suslow,T. and Saltveit,M.

TITLE

Characterization of phenylalanine ammonia-lyase (PAL) gene in wounded lettuce leaf tissue

JOURNAL

2 (bases 1 to 2442)

unpublished

REFERENCE

Campos-Vargas, R., Monogaki, H., Suslow, T. and Saltveit, M.

AUTHORS

Direct Submission

JOURNAL

Submitted (25-AUG-2000) Vegetable Crops, University of California

FEATURES

location/Qualifiers

1..2442

/organism="Lactuca sativa"

/db_xref="taxon:4236"

119..2254

/mole="PAL"

/codon_start=1

/product="phenylalanine ammonia-lyase"

/protein_id="AAU55242.1"

/db_xref="GI:18001007"

/translation="MENGHNHNVNVELIKDPLNNGVAEAATGSHLDEVKMAAEF

RKPVKLSGELTVSVAGIAANDSDIVVELSEARAGVKSMDVMSNKGIDS

YGVITGCGATSHRTKOGALOKELIFLNAIGFNGTETSHLPHSATRAAIYAIN

TLLOGSGIRPELLEATTKFLNNNTPLPACTTASGDLVPLSYIACILGRPSK

ASPTGEVLAKEKAFAGVGGPELQPEGLATVNGTAVGSMASWLPANVIAL

LSEVLSHIREVWOGKPEFDHLTKLHHPGIEAALIEYILDSQVYKAAQYHE

MDLQKPKDRTALRTSPVLDQLEIKSITKMERLNSVNDPLDIVSRNALHG

GNFGRTIGVSMOTFLAIAIGKLMFQPSSELDVFNNGSPNSLGGRNPSLDYF

KGGEIAMAQSELOFLANPVTHVOSAFQHDVNSLGIISARKLAFADVILKMS

TYVALCOSIDLRLHLENKSTVKNVSOVAKVLELNGEPKRPKDLRVYD

REVEFAYLDVQSGVTPMLKRLVLDHALNNGEFLKNTNNTSIPKIAFPPEELKVL

LPRVEGVRIATENDTSLTPRIKATRSYLYLFVVEBELRPSLTGEKVTSPGEEDP

VFIMACKQIILPLELGGNDEPDLIC"

BASE COUNT 693 a 553 c 592 g 604 t

ORIGIN

Query Match 100.0%; Score 2442; DB 8; Length 2442;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCAATCGATCAATACCCATTCACGACAAAGAGTGTAGTCTAGTGTGGAAGT 60
DB 1 GAGCAATCGATCAATACCCATTCACGACAAAGAGTGTAGTCTAGTGTGGAAGT 60
QY 61 ACACAATTGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 120
DB 61 ACACAATTGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 120
QY 121 GGAGAACGCTAATCACTAATGAGTGTAAATGAGTTGTCATCAAGATCCATTGAA 180
DB 121 GGAGAACGCTAATCACTAATGAGTGTAAATGAGTTGTCATCAAGATCCATTGAA 180
QY 181 CTGGGAGCTTGCAGCGGAGCGCTTACCGGAAGTCACCTTATAGGTGAAGAATGCT 240
DB 181 CTGGGAGCTTGCAGCGGAGCGCTTACCGGAAGTCACCTTATAGGTGAAGAATGCT 240
QY 241 TGGGAGCTCAGAAAGCGGCTGTGAAGCTCGAGAGAGAGCGGTTACAGTTTCCAGT 300
DB 241 TGGGAGCTCAGAAAGCGGCTGTGAAGCTCGAGAGAGAGCGGTTACAGTTTCCAGT 300
QY 301 GGGGAGCTCAGAGCTGTGAATGACAGTACCGTGAAGGTGAGTGTGGAAGCGC 360
DB 301 GGGGAGCTCAGAGCTGTGAATGACAGTACCGTGAAGGTGAGTGTGGAAGCGC 360
QY 361 GAGGCTGAGATTAAAGCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 420
DB 361 GAGGCTGAGATTAAAGCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 420
QY 421 TAGTATGCTGACACCGGCTTGCAGCGGCTTGCAGCGGCTTGCAGCGGCTTGCAGCGG 480
DB 421 TAGTATGCTGACACCGGCTTGCAGCGGCTTGCAGCGGCTTGCAGCGGCTTGCAGCGG 480
QY 481 TCGTTTACGAGAGAGCTCATTAGATTTTTCAGACCGCGGAATATTGCGCAATGGAACGA 540
DB 481 TCGTTTACGAGAGAGCTCATTAGATTTTTCAGACCGCGGAATATTGCGCAATGGAACGA 540

QY 541 AACAGGACACACTTGCATTCAGGACAGAGAGCGGATGATGATGATGATGATGATGATGAT 600
DB 541 AACAGGACACACTTGCATTCAGGACACAGAGCGGATGATGATGATGATGATGATGATGAT 600
QY 601 CCGTCTGAGGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 601 CCGTCTGAGGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 TAAZAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 720
DB 661 TAAZAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 720
QY 721 TGTCCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 780
DB 721 TGTCCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 780
QY 781 CCGTCTGAGGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 CCGTCTGAGGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 GTTCTTCAGATTACAGCGGAGAGAGGCTTACGACTTGTACAGGATACGATGATGATGATGAT 900
DB 841 GTTCTTCAGATTACAGCGGAGAGAGGCTTACGACTTGTACAGGATACGATGATGATGATGAT 900
QY 901 CCGGATGCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 901 CCGGATGCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 961 ATCGGATGCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 961 ATCGGATGCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 1021 CAATTAAGACATCAACCGCTGCTCAAAATCGAGCGCGCGCATGATGATGATGATGATGATGAT 1080
DB 1021 CAATTAAGACATCAACCGCTGCTCAAAATCGAGCGCGCGCATGATGATGATGATGATGATGAT 1080
QY 1081 CGGAAGCGATTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1081 CGGAAGCGATTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 1141 AAAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1141 AAAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1201 AATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
DB 1201 AATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
QY 1261 GATGAGCTTCCAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 1320
DB 1261 GATGAGCTTCCAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 1320
QY 1321 AGTTTCATGACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 1380
DB 1321 AGTTTCATGACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 1380
QY 1381 AATTTTCAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
DB 1381 AATTTTCAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
QY 1441 AATTTTCAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
DB 1441 AATTTTCAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
QY 1501 TGTGAGCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
DB 1501 TGTGAGCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
QY 1561 CAATCAAGAGCTTAATTTTCGAGATTATTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
DB 1561 CAATCAAGAGCTTAATTTTCGAGATTATTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620

QY 441 GCTTGGCCGCACTCTTCCAGGAGAACTAAGCAAGCGGCTGTTTACAGAGAGGACGCA 500
 Db 441 GATTGGTCTCTTCTCTAGAGAGAGACTAGAGATGGTGGTCTTACAAAAGAGGCTCA 500
 QY 501 TTAGATTTTGAACCGCGCAATATTCGCAATGCAAGCGAAACAAAGCCACACCTTCAC 560
 Db 501 TTAGATTTTGAACCGCGCAATATTCGCAATGCAAGCGAAACAAAGCCACACCTTCAC 560
 QY 561 ATTGAGCCACAGAGCGGCAATGATGTCAGATTAACACCTCTCTCAGAGGTACTCCG 620
 Db 561 ATTGAGCCACAGAGCGGCAATGATGTCAGATTAACACCTCTCTCAGAGGTACTCCG 620
 QY 621 GCATCGATTTGAGATTTGGAAGCCATCACCAAGTTCTTAAACAAACATCAACCTT 680
 Db 621 GCATCGATTTGAGATTTGGAAGCCATCACCAAGTTCTTAAACAAACATCAACCTT 680
 QY 681 GTTTACCCCTCCGCGTGAACCATCACCCCTCCGCTGACCTTGTCCATTTATCATG 740
 Db 681 GTTTACCCCTCCGCGTGAACCATCACCCCTCCGCTGACCTTGTCCATTTATCATG 740
 QY 741 CCGGCTCTTAACCGGCGCGCCCACTCAAAAGCGTTGGCGCCACCGAGAAAGTCTCA 800
 Db 741 CCGGCTCTTAACCGGCGCGCCCACTCAAAAGCGTTGGCGCCACCGAGAAAGTCTCA 800
 QY 801 ATGCCCAAAAGCGCTTGCAGCGCGAGTTGAAGTGGTGTCTTCAAGTTACAGCGCA 860
 Db 801 ATGCCCAAAAGCGCTTGCAGCGCGAGTTGAAGTGGTGTCTTCAAGTTACAGCGCA 860
 QY 861 AAGAGGGCTAGCACTTGTGAACGGCACCGCGTGGGTCCGGGATGGCTTCCATGTTTC 920
 Db 861 AAGAGGGCTAGCACTTGTGAACGGCACCGCGTGGGTCCGGGATGGCTTCCATGTTTC 920
 QY 921 TATTGATGCTATGACTTGGTGTGTGGAAGTGTATTCGGCATTTCTCGTAGG 980
 Db 921 TATTGATGCTATGACTTGGTGTGTGGAAGTGTATTCGGCATTTCTCGTAGG 980
 QY 981 TTATGCAAGGAAAGCGCGATTTACCGATCTGACACACAAATGGAAGCATCCCTG 1040
 Db 981 TTATGCAAGGAAAGCGCGATTTACCGATCTGACACACAAATGGAAGCATCCCTG 1040
 QY 1041 GTCAAAATGAGAGCGGCGGATGATGAGTATTTTGGAGGAGAAAGCATACCTGCAAG 1100
 Db 1041 GTCAAAATGAGAGCGGCGGATGATGAGTATTTTGGAGGAGAAAGCATACCTGCAAG 1100
 QY 1101 CCGCGCAAAAGGTTCACGAAATGACCGGTTACAGAAACCAAAACAGATGTTATGTC 1160
 Db 1101 CCGCGCAAAAGGTTCACGAAATGACCGGTTACAGAAACCAAAACAGATGTTATGTC 1160
 QY 1161 TCCGTACATCTGCCCAATGGCTGCGACCTGCAATGGAAGTATCGCATCAACCAAAA 1220
 Db 1161 TCCGTACATCTGCCCAATGGCTGCGACCTGCAATGGAAGTATCGCATCAACCAAAA 1220
 QY 1221 TGAATGAGAGGAGATCAACTCCGTTAACGACAAACCCCTATGACAGCTGTGCAAAA 1280
 Db 1221 TGAATGAGAGGAGATCAACTCCGTTAACGACAAACCCCTATGACAGCTGTGCAAAA 1280
 QY 1281 AAGCTTTACAGCGTGTAACTTCAAGGAACCCCAATGGAAGTTTCATGAGCAACACC 1340
 Db 1281 AAGCTTTACAGCGTGTAACTTCAAGGAACCCCAATGGAAGTTTCATGAGCAACACC 1340
 QY 1341 GTTCGCCCATTTGCAATCGGAAACCTGATGTTGCTCAATTTTCTGAGCTGTTAAGC 1400
 Db 1341 GTTCGCCCATTTGCAATCGGAAACCTGATGTTGCTCAATTTTCTGAGCTGTTAAGC 1400
 QY 1401 ATTCTACAAACATGATTAACATCAATCTCTCGGTGGAGAGTAACCTTGTGTTGACT 1460
 Db 1401 ATTCTACAAACATGATTAACATCAATCTCTCGGTGGAGAGTAACCTTGTGTTGACT 1460
 QY 1461 ACGGTTCAAGGTGAGAAATGCGCATGCTTCTTACTGTTGAGCTTCAAGTTCTCG 1520
 Db 1461 ACGGTTCAAGGTGAGAAATGCGCATGCTTCTTACTGTTGAGCTTCAAGTTCTCG 1520

QY 1521 CAATTCACATCCACCAACCATGTTCAAAAGCGCGCAACAAACAAATCAAGAGCTTAAATTC 1580
 Db 1521 CAATTCACATCCACCAACCATGTTCAAAAGCGCGCGCAACAAACAAATCAAGAGCTTAAATTC 1580
 QY 1581 TCGATTTAATTTACGAGGAGAAACCGTAGAAGTATGATCTTAAATCTATGCTG 1640
 Db 1581 TCGATTTAATTTACGAGGAGAAACCGTAGAAGTATGATCTTAAATCTATGCTG 1640
 QY 1641 CGACATCTTACTGCTCTATGCCAATTCATGATGATTTAAGCATTTTGAAGAGACATGA 1700
 Db 1641 CGACATCTTACTGCTCTATGCCAATTCATGATGATTTAAGCATTTTGAAGAGACATGA 1700
 QY 1701 AATCGACATGTAAGAACAACCGTAAGCCAGTCCGCAAAAGAGTCTTCAACCATGAGCGCTCA 1760
 Db 1701 AATCGACATGTAAGAACAACCGTAAGCCAGTCCGCAAAAGAGTCTTCAACCATGAGCGCTCA 1760
 QY 1761 ACGGAGAGCTCCACCGCTGAGATTTCTCGGAAAGATTTCTTCTGTTTGTATGTTG 1820
 Db 1761 ACGGAGAGCTCCACCGCTGAGATTTCTCGGAAAGATTTCTTCTGTTTGTATGTTG 1820
 QY 1821 AATGATCTTGTCTTACATTCGACGAGCTTTGACGAGGACATACCATTAATGAGAAAG 1880
 Db 1821 AATGATCTTGTCTTACATTCGACGAGCTTTGACGAGGACATACCATTAATGAGAAAG 1880
 QY 1881 TCCGACAGGTTCTGCTGACACACCGCTTAAACCAACGCGGAAACGAGAGACATTAACA 1940
 Db 1881 TCCGACAGGTTCTGCTGACACACCGCTTAAACCAACGCGGAAACGAGAGACATTAACA 1940
 QY 1941 CCTCATCTTCTTCAAAAGATCTTACTTCTGAAAGAAATTTAAAGTCTTCTTCTGAAAG 2000
 Db 1941 CCTCATCTTCTTCAAAAGATCTTACTTCTGAAAGAAATTTAAAGTCTTCTTCTGAAAG 2000
 QY 2001 AAGTTGAAGGTGTGTAAGATTCGCTTATGAAATGATTAATTCATTCACCAACAGATTA 2060
 Db 2001 AAGTTGAAGGTGTGTAAGATTCGCTTATGAAATGATTAATTCATTCACCAACAGATTA 2060
 QY 2061 AAGCTTGCAGATCTGACCGCTTGTATAGCTTGTGTAAGGAGACAGCTGCGCAGAGGTTT 2120
 Db 2061 AAGCTTGCAGATCTGACCGCTTGTATAGCTTGTGTAAGGAGACAGCTGCGCAGAGGTTT 2120
 QY 2121 TGAACGAGAAAGGTGAGCGTGGCGGAGAGAGGATTCGACAGGAGTTCACGCGCATGT 2180
 Db 2121 TGAACGAGAAAGGTGAGCGTGGCGGAGAGAGGATTCGACAGGAGTTCACGCGCATGT 2180
 QY 2181 GCAAAAGTCAATTTATGATCCGTTTGAAGTGTCTTGAAGGAGTGAATGAGGAACTTC 2240
 Db 2181 GCAAAAGTCAATTTATGATCCGTTTGAAGTGTCTTGAAGGAGTGAATGAGGAACTTC 2240
 QY 2241 TTCCATATATGTAAGAAAGTGAAGTGAACCGTTTGAATTTGTATATATCTGTT 2300
 Db 2241 TTCCATATATGTAAGAAAGTGAAGTGAACCGTTTGAATTTGTATATATCTGTT 2300
 QY 2301 TTTTGTGTTT 2309
 Db 2301 TTTTGTGTTT 2308

RESULT 3
 AF401636
 LOCUS AF401636 2444 bp mRNA linear PLN 08-AUG-2001
 DEFINITION Rehmannia glutinosa phenylalanine ammonia-lyase (PAL1) mRNA,
 complete cds.
 ACCESSION AF401636
 VERSION AF401636.1 GI:15100058
 KEYWORDS
 SOURCE Rehmannia glutinosa.
 ORGANISM Rehmannia glutinosa.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Lamiales; Lamiales incertae sedis;
 Rehmannia.
 REFERENCE 1 (bases 1 to 2444)
 Yun,S.-J., Lee,B.-K., Park,M.-R., Srinivas,B. and Chun,J.-C.

Db 1029 GAACACATTTGGATGCTACCTCTTAAAGGCTTCTGCAAGAGGCAACATGCAAT 1088
 QY 1127 GAGTATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 1186
 Db 1089 GCTTATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 1148
 QY 1187 GCTTATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 1246
 Db 1149 GCTTATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 1208
 QY 1247 AAGCAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 1306
 Db 1209 AAGCAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 1268
 QY 1307 GAGCAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 1366
 Db 1269 GAGCAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 1328
 QY 1367 GCTTATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 1426
 Db 1329 GCTTATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 1388
 QY 1427 AATCTTATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 1486
 Db 1389 AATCTTATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 1448
 QY 1487 AATCTTATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 1546
 Db 1449 AATCTTATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 1508
 QY 1547 AGCCGCGCAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 1606
 Db 1509 AGCCGCGCAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 1568
 QY 1607 GCAACAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 1666
 Db 1569 GCAACAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 1628
 QY 1667 GCTTATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 1726
 Db 1629 GCTTATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 1688
 QY 1727 CAATGCGCAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 1786
 Db 1689 CAATGCGCAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 1748
 QY 1787 TGGCAAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 1846
 Db 1749 TGGCAAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 1808
 QY 1847 GTTATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 1906
 Db 1809 GTTATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 1868
 QY 1907 CTAACAAACCGCAACCGCAACCGCAACCGCAACCGCAACCGCAACCGCAACCGCAAC 1966
 Db 1869 CTAACAAACCGCAACCGCAACCGCAACCGCAACCGCAACCGCAACCGCAACCGCAAC 1928
 QY 1967 TTTCAAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 2026
 Db 1929 TTTCAAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 1988
 QY 2027 GAGCAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 2086
 Db 1989 GAGCAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 2048
 QY 2087 AGCTTATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 2146
 Db 2049 AGCTTATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 2108
 QY 2147 GAGCAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 2206
 Db 2109 GAGCAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 2168

QY 2207 TTTGAGTCTCTTGGAGGCTGGAAATGGGGAACCTCTTCCAAATATGTATAGCAAGCTAGCT 2266
 Db 2169 TTTGAGTCTCTTGGAGGCTGGAAATGGGGAACCTCTTCCAAATATGTATAGCAAGCTAGCT 2228
 QY 2267 GAAACCTTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2326
 Db 2229 TTTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2288

QY 2327 TG 2328

Db 2329 TG 2330

RESULT 5

PCPBL3

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

ATTNRS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

ORIGIN

BASH COUNT

Query Match

Host local Similarity

Matches 1634

Conservative 0

Score 177.8

Pos 9

Length 2403

Prod. No. 130293

Indels 6

Gaps 2

Indels 6

Gaps 2

135 ACCTTAATGAGTCGTTAAATGAGTTGTGATCAACAGATCCATGTAAGTGGGAGTTGGAG 194
126 AGGGAGAGGAGATTAGATTTGTGCATGAAAGAAAGAGATCTTTAACTGGGGAGTGGCTG 185
195 CGAGAGCGTTGACCGGAATGACCTTGTGAGGTGAGAGAGATGTTGCCGAGTTTCAGAA 254
186 CCGAGGCAATTCGACGAGTCTATTTGGACGAAGTTAAAGAGATGTGGCTGATACAGGA 245
255 AGCGGTGTGAAAGTGTGAGAGAGAGAGCTTACAGTTTCTAGGTGAGGAGATCCAG 314
246 ACCCGGTGTGAGGTGAG 305
315 CTGCTAATGAGTACACACCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 374
306 CTAGGATATATAGT---GTGTTAAGTGTGAGCTTCCGAGAGAGAGAGAGAGAGAGAG 362
375 AGCGGAGTATGATGTTGGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 434
363 AGCGTATGATGAGTGTGAG 422
435 CCACGAGTGTGAG 494
423 CTACTGCTTGTGCTGCTACTTCTCATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 482
495 AGCTGATTAATTTTGAACCGCGGAGATTTCCGCAATGAGAGAGAGAGAGAGAGAGAGAG 551
483 AGCTATTAGTGTGTTGAATGCTGATTAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAG 542
552 CAGTTGACATTTGAG 611
543 CATTCACAGCTCCGCAACAG 602
612 GTTATCCGCGATCCGATTCGAGATCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 671
603 GCAATTACAGAGATCCGATTCGAGATCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 662
672 TCACCCCTGTTTACCCCTCCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 731
663 TTACTGCTGTTGAG 724
732 CATACATCCGCGCTTGAACCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 791
723 CCTACATTCGAGATCTTCTCACTGCTGCTCCCACTCCAGAGAGAGAGAGAGAGAGAGAG 782
792 AAGTCTCTAATGATGAG 851
783 TTAACATCTGAGCGCGAG 842
852 TACAGCGGAG 911
843 TACAGCGGAG 902
912 CCATGCTCTAATTTGATCTTATGATCTTATGATCTTATGATCTTATGATCTTATGATCT 971
903 CTATGCTCTAATTTGATCTTATGATCTTATGATCTTATGATCTTATGATCTTATGATCT 962
972 TGGCTGAGAGTATGATGAG 1031
963 TGGCTGAGAGTATGATGAG 1022
1032 ATACACCTGAGTCAATTCAG 1091
1023 ACCATCTCTGAGTCAATTTAGAGTGTGATTAATGAGAGAGAGAGAGAGAGAGAGAGAG 1082
1092 AGCTGAG 1151
1083 AGCTTAAAGCTGTGCAAAAAGTACATGATGATGATGATGATGATGATGATGATGATGAT 1142
1152 GTTATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1211
1143 GATATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1202

1212 CAACAAATGATCCAG 1271
1203 CAACATAATGATGAG 1262
1272 CCGAAG 1331
1263 CAG 1322
1332 ACAG 1391
1323 ACAG 1382
1392 TGGTAAAG 1451
1383 TGTGATCAG 1442
1452 GTTGAAGTACAG 1511
1443 GCTTGAATTTAGATTCAG 1502
1512 AGTTTCTGAG 1571
1503 AGTTTCTGAG 1562
1572 TTAATTCCTGAG 1631
1563 TGAATCTGAG 1622
1632 TCAATGCTGAG 1691
1623 TCAATGCTGAG 1682
1692 AGAGATGAG 1751
1683 AAGATTTGAG 1742
1752 TGGGCTGAG 1811
1743 TGGGCTGAG 1802
1812 TGGGCTGAG 1871
1803 TGGGCTGAG 1862
1872 TGGGCTGAG 1931
1863 TGGGCTGAG 1922
1932 ACAGTAAAG 1991
1923 ACTTGAAG 1982
1992 TACCGAAG 2051
1983 TCCCTAAAG 2042
2052 ACAGTAAAG 2111
2043 ACAGTAAAG 2102
2112 GAGGCTGAG 2171
2103 CTGAGTATCTTACAG 2162
2172 CCGGAGTGTGAG 2231
2163 CAGCAATGTGAG 2222
2232 GGGAG 2291
2223 GTGCTGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2282
2292 TATCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2333

QY	1427	AATCTCTCCGTTGACAGTAACTCACTATTGTGACATCCGGTTCACAAAGTCCAGAAATGCC	1486
Db	1418	AATCTCTCCGTTGACAGTAACTCACTATTGTGACATCCGGTTCACAAAGTCCAGAAATGCC	1477
QY	1487	ATGAGTTCTTAATGTTCTGATGATTAATTTCTGCAATATGATATGATATGTTCAA	1546
Db	1478	ATGAGTTCTTAATGTTCTGATGATTAATTTCTGCAATATGATATGATATGTTCAA	1537
QY	1547	AGGCGCGAATCAACAAATCAAGAGCTTAATCTCTCGATTAATTTCAAGAGAGAAAC	1606
Db	1538	AGGCGCGAATCAACAAATCAAGAGCTTAATCTCTCGATTAATTTCAAGAGAGAAAC	1597
QY	1607	GCAGAAACCATCTGCAATCTTTAAACTCATCTGCTGACATACCTTAGTCGCTCAATGCCA	1666
Db	1598	GTCGAGGCTCTGCATCTTAACATCATGTCGTGACCTACTTGGTGGCTTATGCAAC	1657
QY	1667	TCCATGCTATTACGCCATTTTGAGAGAAATCATGAATGACATGAGAAACACCGTTAAG	1726
Db	1658	GCCCTCATCTGAGGGATGTGAGAGAGAAATCAAGAGCTGCCCTGCAAAAACACATGTCAG	1717
QY	1727	CAATGCTCGCAAAAAGCTGCTCACTGAGGCTTCAACGCGCTGAGCTCCACCGTGCAGATTC	1786
Db	1718	CAAGTTCCGCAAGGAACTCTCAATTTGGCTGCTAATGGCAATCCACATCCATCGATTC	1777
QY	1787	TCCGAGAAAGATCTCCCTCCGTTGTGATGCTGATACGCTTTCGCTTACATGACAGAC	1846
Db	1778	TGCGAGAAAGATCTCCCTCCGTTGTGATGCTGATACGCTTTCGCTTACATGACAGAC	1837
QY	1847	GTTTGCAGGCGGCATATCCCATTAATGACAAAGCTCCGACAGGTTCTGTGACACAGCT	1906
Db	1838	CCCTGCTCCCGACCTATCCCGTTGATGAGAAAGCTGAGCGAGCTGCTGTGACACAGCT	1897
QY	1907	CTAATCAAGGCGAAGAGGAGAAACCTTACCTCATCTTCCCAAAAGATGCGTACC	1966
Db	1898	CTGAGAAAGCGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1957
QY	1967	TTCGAGAGAGATTTGAAAGTCTCTGTTACCGAAAGAGTTGAAGCTGTTAGATGCTTA	2026
Db	1958	TTCGAGAGAGAGTGTGAAGGCTCTGCTACCCAAAGAGGCTGAGAGCGCAAGATGCGCTC	2017
QY	2027	GAGGATATATCATTTGTGATTTCCAAACAGGATTAAGTTTCACATGATGATCCGTTGAT	2086
Db	2018	GAGGATGCTGATGCTGCTGCTGCGCAACCGCATAGCGAATCCCGCTCTCTCTCTCTTA	2077
QY	2087	AGGTTTGTAAAGGAGAGAGCTGCGAGAGAGGTTTGTGACGCGAGAAAGAGTGAAGTGCAG	2146
Db	2078	AGGTTTGTAAAGGAGAGAGCTGCGAGAGAGGTTTGTGACGCGAGAAAGAGTGAAGTGCAG	2137
QY	2147	GGAGAGAGCTTGCACAGGCTGTTCAACCGCATGTCGAAAGCTCAAAATTAATTCCTTC	2206
Db	2138	GGAGAGAGAGTGCACAGAGGCTGTTCAACCGCATGTCGAAAGCTCAAAATTAATTCCTTC	2197
QY	2207	TTGAGAGGTTTGTGAGAGGTTGAAATGAAATGTTTCAATTAATGTTT	2254
Db	2198	TTGAGAGGTTTGTGAGAGGTTGAAATGAAATGTTTCAATTAATGTTT	2245

RESULT 7	
NTPHEAL	
LOCUS	NTPHEAL 2462 bp mRNA linear PLN 16-JAN-1995
DEFINITION	X.tabacum (Samsun NN) mRNA for phenylalanine ammonia-lyase.
ACCESSION	X78269
VERSION	X78269.1 GI:633596
KEYWORDS	phenylalanine ammonia-lyase,
SOURCE	Nicotiana tabacum.
ORGANISM	Nicotiana tabacum
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asterales; euasterids I; Solanales; Solanaceae; Nicotiana. 1 (bases 1 to 2462)
REFERENCE	Pellierin,L., Rohlfisch,O., Fritli,B. and Legrand,M.
AUTHORS	phenylalanine ammonia-lyase in tobacco. Molecular cloning and gene expression during the hypersensitive reaction to tobacco mosaic
TITLE	

JOURNAL	PLANT and the response to a fungal elicitor
MEDLINE	Plant Physiol. 106 (3), 877-886 (1994)
PUBLISHED	95125127
REFERENCE	7824656
AUTHORS	2 (bases 1 to 2462)
JOURNAL	Pellegrini, L.
FEATURES	Direct Submission
source	Submitted (17-MAR-1994) L. Pellegrini, Institut de Biologie Molec.
	des Plantes, 12 rue du General Zimmer, 67084 Strasbourg, FRANCE
	Location/Qualifiers
	1..2462
	/organism="Nicotiana tabacum"
	/cultivar="Samsun NN"
	/db.xref="taxon:4097"
	/clone="PAL.E"
	/tissue.type="leaves"
	/clone_id="lambda ZAP II vector, cDNA from tobacco leaves
	72 hrs after infection by Tobacco Mosaic Virus (TMV)"
	/dev_stage="30 day old"
	122..2260
	/EC_number="4.3.1.5"
	/codon_start=1
	/product="phenylalanine ammonia-lyase"
	/protein_id="CA55075.1"
	/db.xref="GI:633597"
	/db_xref="SWISS-PROT:P45733"
	/translation="MAGVAQNGHQEDPCVKYDPLNWEKMAADSLKSHLDEVKKMVAE
	EKKPVYKLGTELLVAVQVAAIAAKDNATKYVELSEGAAGYKASDWMDSKGTD
	SGVGTGFGATSHRRKNGALOKELIRELNVGNGSGTESCHTLPQSTPAALVRI
	NLLDQSGIRPELLEAITKLKLNHYETPCPLRGITLASGDIPLVSTIAGLITGRPS
	KVIGNGPILNAEPAPVACVNSGFPELOPKREGIALVNGTAGSGLASMTIFDANILA
	VSEYLSKIFAEVWNGKPERFDLHTRKLNHFOUIBAAMKPHLIDDSVYKAPOKIH
	ETDLPQKPDORVALRTSPQWQLIEVLRSAFTKMIEREINSDNPLIDVSRKALH
	GNFQETIIVGSDMNAFLASIGKLMFQSEFELVYNNGLPSNLTAGENPSLDGS
	FGSGIIMASYCSELQELANPVTHYQSEQHNQVNSIGLISAKTIEADVILDKLS
	SYVALCOALDILRLLEENLRNAKMTVSOVAKRTITGMANGELHPSPFCEDLLRV
	DEEYVLRADDACSNAPLMOKLROYVLVHAIENLQENPKNKANSSIPQKILAFEEELKA
	VIKPKVESARISLENGCPAIANRIKCDREYPLKRYREELCALELLTGKNSRQEBED
	KYFTMCMGQITIDLELCELEKNNGAFLPIC"

Spermatophyta; Magnoliophyta; eudicotyledons, core eudicots, Asteridae; Ericales; Theaceae; Camellia.

REFERENCE
AUTHORS
TITLE
JOURNAL
2 (bases 1 to 2330)
Matsumoto, S.
TITLE
Direct Submission
Submitted (22-JAN-1994) Satoru Matsumoto, Natl. Res. Inst. of Vegetables, Ornamental Plants and Tea; Kanaya 2769, Kanaya, Shizuoka 428, Japan (Tel:0547-45-4101, Fax:0547-46-2169)
FEATURES
source
1..2330
/organism="Camellia sinensis"
/cultivar="Yadukita"
/db_xref="taxon:4442"
/clone="A21"
/issue_type="Young leaf"
/clone_lib="lambda gll1"
132..2276
/codon_start=1
/product="phenylalanine ammonia-lyase"
/protein_id="BA05643.1"
/db_xref="gi:662271"
/translation="MDSTAIANGVSGSGSPGCLKDPNLNMGVAEAMKSHLEEVKG
MVEERKPVRLGEGELLISQVLAIVRGSEVAVESRECVKSNMVMKSNMKG
TDSYGTTCGATSHRRKTKGLOKELIPLNAGFNGTRESCHLPOSATPAALY
RINTLQSGSTREPLELAIKSLNNITPCLPRTGTTASGDLVPLSTAGLITGRH
NSKAVPTGELIHPKEARLAVGEGFPEIOPRGLAVGTAVSGGLASMTLFEANI
LAVSEVLSIAFAEVMQKPEFDLHKHKGQLEAAIMHILIDSSTYKAAOK
LHENDLPKQDRYALSTPQWGLPIEIVRSSTSEIENSVDNPLINSRKA
LHGNEOSTPIGSDMTPLAVASIGLMPSESELYNDVNGLSNLSGPNPLD
YGFKAELIAMAAYCSLOPLANPVTNHVGAHODNVDNLGISSKTAFAVDILKL
MSTYVALCOADVLRHPEENLHTVNTSYSAKRYLMVNGCELHPKRCCKDLK
VVDREYIPATIDPDSATYPLMKRLQVLRVLAHAKNGSKNSLSTIPKIRAFEEI
KTLPEVESTRADIEGNSAIPRIIECHSYLPKRVREELTELLTEKVSPEE
EDKVFALCKGEMIDPLMDCKEMNGAPLIC"

polyA_signal
polyA_site 2330
BASE COUNT 630 a 496 c 595 g 609 t
ORIGIN

Query Match 51.4%; Score 1254.2; DB 8; Length 2330.
Best local similarity 74.7%; Pred. No. 1e-289;
Matches 1590; Conservative 0; Mismatches 533; Indels 6; Gaps 1,

QY 158 TTGTGCATCAAGCATTCATCAAGTGGAGAGTTCGACGCGGAGAGTGCAC 217
DB 186 TTTTGTCTAAGAACCTTTGATTTGGAGAGTGGCGGAGAGCAATGAAGGAGTCTAT 245
QY 218 CTTCATGAGTCAAGACAGATGTTCCGAGCTTCAGAAAGCCGCTGCGGAGCA 277
DB 246 TTGAGAGAGTCAAGCGCATGCTGGAGAGCTTCGGAAGCCAGTGGTAGCTGGAGGG 305
QY 278 GAGAGCTTAAAGATTCTTCAAGTGGAGAGTGGAGAGTGGAGAGTGGAGAGTGG 337
DB 306 GAGAGCTTAAAGATTCTTCAAGTGGAGAGTGGAGAGTGGAGAGTGGAGAGTGG 359
QY 338 AAGGTGAGCTTCGAGAGAGCCGAGAGGCTGAGTGAAGGCGAGTGAAGTGGAGTGG 397
DB 360 GCGGTGAGCTTCGAGAGTGGAGAGGAGAGGAGAGTGAAGGCGAGTGAAGTGGAGTGG 419
QY 398 GAGAGCATTAAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 457
DB 420 GAGAGTGAAGTGAAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 479
QY 458 GAGAGTGAAGTGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 517
DB 480 CATAGAGAGAGTGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 539

QY 518 GGAATATTGGCAGATTCAGAGGAG 577
DB 540 GGAATATTGGCAGATTCAG 599
QY 578 GGCATGATTCGAG 637
DB 600 GGCATGATTCGAG 659
QY 638 TTGGAAGGAG 697
DB 660 TTGGAAGGAG 719
QY 698 ACCATGACGAG 757
DB 720 ACTATGACGAG 799
QY 758 GCGCCAG 817
DB 780 GCGCCAG 839
QY 818 GCTGAG 877
DB 840 GCTGAG 899
QY 878 GTTAAAG 937
DB 900 GTTAAAG 959
QY 938 GTTAAAG 997
DB 960 GTTAAAG 1019
QY 998 GAGTTTACGAG 1057
DB 1020 GAGTTTACGAG 1079
QY 1058 GCGATGATGAGAGATATTGAG 1117
DB 1080 GCGATGATGAGAGATATTGAG 1139
QY 1118 GAAATGAG 1177
DB 1140 GAAATGAG 1199
QY 1178 TGGCTGAG 1237
DB 1200 TGGCTGAG 1259
QY 1238 AATTCGATGAG 1297
DB 1260 AATTCGATGAG 1319
QY 1298 AATTCGATGAG 1357
DB 1320 AATTCGATGAG 1379
QY 1358 ATGCGAG 1417
DB 1380 ATGCGAG 1439
QY 1418 TTACATGAG 1477
DB 1440 TTACATGAG 1499
QY 1478 GAAATGAG 1537
DB 1500 GAAATGAG 1559
QY 1538 CATGTTTAAAG 1597
DB 1560 CATGTTTAAAG 1619
QY 1598 AGGAAAG 1657

[illegible]

COMMENT	FEATURES	SOURCE
<p> GenBank: U00096.1 (1997) This sequence version replaced gi:464217. Sequence updated (22-Oct-1997). location/Qualifiers 1..2527 </p>	<p> /organism="Nicotiana tabacum" /db_xref="Taxon:4097" /cell_line="Bright Yellow T-19" 1..2527 /gene="TOBPAL1" 205..2143 /gene="TOBPAL1" /EC_number="4.3.1.5" /codon_start=1 /evidence=experimental /product="phenylalanine ammonia-lyase" /protein_id="AA022963.1" /db_xref="GI:2570156" </p>	<p> CDS 205..2143 /gene="TOBPAL1" /EC_number="4.3.1.5" /codon_start=1 /evidence=experimental /product="phenylalanine ammonia-lyase" /protein_id="AA022963.1" /db_xref="GI:2570156" </p>
<p> Query Match Host local Similarity 71.8%; Score 1250.8; DB 8; Length 2527; Matches 1656; Conservative 0; Pred. No. 6.96-289; Mismatches 647; Indels 1; Gaps 1; </p>	<p> polyA_signal polyA_site BASE COUNT 741 a 517 c 557 g 712 t ORIGIN </p>	<p> 132 133 TCACGCTAAACGAGCGCTTAATGAGCTTACGATCAAGCAACGACGAGCTTCC 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560</p>

Db 1116 GCTCTGAGACATCTCCACAGTGGCTGGCCCTCAGATTGAGTATACCAAGCAAC 1175
 QY 1217 AAAATGATCGAGAGGAAATCAATTCCGTACACACACCACTTATGACAGTTTCAGA 1276
 Db 1176 AAAATGATCGAGAGGAAATCAATTCCGTACACACCACTTATGACAGTTTCAGA 1235
 QY 1277 AACAAAGCTTACAGCGGTGTAACCTTCAAGGAGCAACCCCAATCCGAGTTTCAGACAC 1336
 Db 1236 AACAAAGCTTACAGCGGTGTAACCTTCAAGGAGCAACCCCAATCCGAGTTTCAGACAC 1295
 QY 1337 ACCGCTCTGCGGATTTGTCATGAGAAATCAATGTTGCTCAATTTTGAAGTGGTT 1396
 Db 1296 TCTACATTAGCGCTTGGCTTCATGAGAAATCAATGTTGCTCAATTTTGAAGTGGTT 1355
 QY 1397 AACGATTTCTACAAATGATGATACCTGATGATGCTGCTGATGATGATGATGATGAT 1456
 Db 1356 AACGATTTCTACAAATGATGATACCTGATGATGCTGCTGATGATGATGATGATGAT 1415
 QY 1457 GACTAGCGGCTTCAAGGTGAGAAATGCGCATGGCTTCTACTGCTGAGCTTCAAGTTT 1516
 Db 1416 GATTATGATTTCAAGGCGGTGAGAAATGCGCATGGCTTCTACTGCTGAGCTTCAATTC 1475
 QY 1517 CTGCGAAATCGAGTCAACCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1576
 Db 1476 TTGGCTAATCCGCTCAACCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1535
 QY 1577 TCTCTGAGTATTAATTCAG 1636
 Db 1536 TCCCTGGGTTGATCTCTGAG 1595
 QY 1637 TCGTCGACATTAATGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 1696
 Db 1596 TCCCTGACATTAATGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 1655
 QY 1697 ATGAATGACAGTGAAG 1756
 Db 1656 TTGAAG 1715
 QY 1757 GTCAAG 1816
 Db 1716 GTCAAG 1775
 QY 1817 GGTGATGACGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1876
 Db 1776 GCGGATGACGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1835
 QY 1877 AAGTCCGACAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1936
 Db 1836 AAGTCCGACAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1895
 QY 1937 AATACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1996
 Db 1896 AATACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1955
 QY 1997 AAGCAAGTGAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2056
 Db 1956 AAGCAAGTGAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2015
 QY 2057 ATTAAAGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2116
 Db 2016 ATTAAAGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2075
 QY 2117 TTTTGGAG 2176
 Db 2076 TTTTGGAG 2135
 QY 2177 ATGTCGAAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2236
 Db 2136 ATGTCGAAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2195
 QY 2237 CCTTCGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2296
 Db 2196 CCTTCGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2255

QY 2297 TGTTTT 2303
 Db 2256 GGGTGT 2262

RESULT 14
 PCPAL2
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

PCPAL2
 P. crispum mRNA for phenylalanine ammonia-lyase 2.
 X81158
 81158.1
 GI:534892
 deaminase; PAL-2 gene; phenylalanine ammonia-lyase.
 Petroselinum crispum
 Petroselinum crispum
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids II; Apiales; Apiaceae; Petroselinum.
 1 (bases 1 to 2384)
 Appert, C., Logemann, E., Hahlbrock, K., Schmid, J., and Amrhein, N.
 Structural and catalytic properties of the four phenylalanine
 ammonia-lyase isoenzymes from parsley (Petroselinum crispum Nym.)
 Eur. J. Biochem. 225 (1), 491-499 (1994)

JOURNAL
 MEDLINE
 PUBMED
 PUBLISHED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

95010141
 7925471
 2 (bases 1 to 2384)
 Appert, C.
 Direct Submission
 Submitted (22-AUG-1994) C. Appert, Institut fuer
 Pflanzenwissenschaften, Eidgenossische Technische Hochschule,
 Zuerich, Universitaetsstrasse 2, 8092 Zurich, SWITZERLAND
 location/Qualifiers
 1. .2384
 /organism="Petroselinum crispum"
 /db_xref="taxon:4043"
 /note="tetramere subunit"
 59. .2209
 /gene="PAL2"
 59. .2209
 /EC number="4.3.1.5"
 /note="deaminase subunit"
 /codon_start=1
 /product="phenylalanine ammonia-lyase 2"
 /protein_id="CAA57056.1"
 /db_xref="GI:534893"
 /translation="MENGCAITNCHVNGCMDFQMKTEDEPLVYMGIAEAMTGSHTDE
 VKRMVAEYRPPVKLGCEPLTTSOVAIAISARGSGVIVLEISPAARAGYKASIDVYMS
 MKGTDSEYGVTFQATSHRTRKOGALOKELIRPLNAGIRGNSDNLTPHSATRAAM
 LVRIINLUGTSIREILIEATITKPLNWIPLDPLRGTITRSGDVLPLSTVAGLLTG
 RPSKAVGPPTGLISPEEAFKLAGVEGFEFLQPEKGLAVNGTAVGSGMSVLFEEA
 NIILAVIAEWSAIFAEVMOGKEPETHLHKLHPGGIEAAAIWEHILDSAVYKKA
 OKLHEMDPLQKPRODRLRTSPQIIEVIRSTKMIEREIINSVDNPLIIVSRN
 KAIHGNFOGTPIGMSMDNTRLIAIICKLMAQPSLVIENPYNNGLPNSLGGNPS
 LDYCFRGAETIAMSVCSELOPLANTVNTVOSARQNDVNSICILSRKTSPEVILL
 KLMSTPFLVGLCAILDLRHLEBENKSTVANTVSYAKKVLIMVNGELHRSRFEKDL
 LRFVDEYIFATYIDDPKSAFVPLDMKLTROTLVEHMLKMDNEPNNNTITFQIATFD
 ELKALLPEKVESRALLESQNPATPNP:EECHSVPIYFVKRELIEVLTGEKVTSPG
 EEPDKVFIAMSKELIDPLECLSNMGALPDIQ"

BASE COUNT 700 a 427 c 581 g 676 t
 ORIGIN

Query Match 50.38; Score 1227.8; DB 8; Length 2384;
 Best Local Similarity 72.14; Pred. No. 228.283;
 Matches 1661; Conservative 0; Mismatches 622; Indels 20; Gaps 4;

QY 135 ACGTTAATGAGCGGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 194
 Db 99 ACGGAAACGAGATGATTTGATGAAGACCGAAGATCCGATGATGAGGAAATCGCGG 158
 QY 195 CGAGCGCTTGACCGGAGATCACCTTGATGAGTGAAGAGATGATGCGGAGATTGACAA 254
 Db 195 CGAGCGCTTGACCGGAGATCACCTTGATGAGTGAAGAGATGATGCGGAGATTGACAA 254

Db 159 CCGAGCCATGACAGGAGTCATTTGCAATGCAAGCAAGATGCTGGCTGATTAAGCA 218
QY 255 AGCTGGTGGTGAAGTCTGAGAGAGAGAGTATAGTTTCTAGAGTGGTGGAGTGGAG 314
Db 219 AGCTGGTGGTGAAGTCTGAGAGAGAGAGTATAGTTTCTAGAGTGGTGGAGTGGAG 278
QY 315 CTCTGATGACAGTGGAGCCGTCAGAGCTGGAGAGCCGTCAGAGCCGTCAGAGTGA 374
Db 279 CTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 335
QY 375 AGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 434
Db 336 AGGCTAGTACAGTGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 395
QY 435 CTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 494
Db 396 CTAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 455
QY 495 AGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 554
Db 456 AGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 509
QY 555 TTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 614
Db 510 TACGACAGTGGAG 569
QY 615 AGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 674
Db 570 AGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 629
QY 675 GCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 734
Db 630 GCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 689
QY 735 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 794
Db 690 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 749
QY 795 TCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 854
Db 750 TCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 809
QY 855 AGCGGAG 914
Db 810 AGCGGAG 869
QY 915 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 974
Db 870 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 929
QY 975 CTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1034
Db 930 CTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 989
QY 1035 AGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1094
Db 990 AGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1049
QY 1095 TCAAG 1154
Db 1050 TCAAG 1109
QY 1155 ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1214
Db 1110 ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1169
QY 1215 CCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1274
Db 1170 CCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1229
QY 1275 GAAAG 1334
Db 1230 GAAAG 1289

QY 1335 ACAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1394
Db 1290 ATAG 1349
QY 1395 TTACAGATGCTGCAAG 1454
Db 1350 TTACAGATGCTGCAAG 1409
QY 1455 TGAAGATGCTGCAAG 1514
Db 1410 TGAAGATGCTGCAAG 1469
QY 1515 TTGCTGCAAGATGCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1574
Db 1470 TTGCTGCAAGATGCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1529
QY 1575 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1634
Db 1530 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1589
QY 1635 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1694
Db 1590 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1649
QY 1695 ACATGAGATGCTGCAAG 1754
Db 1650 ACATGAGATGCTGCAAG 1709
QY 1755 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1814
Db 1710 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1769
QY 1815 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1874
Db 1770 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1829
QY 1875 AGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1934
Db 1840 AGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1889
QY 1935 CTAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1994
Db 1890 CTAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1949
QY 1995 GGAAG 2054
Db 1950 GGAAG 2009
QY 2055 GCAATGAG 2114
Db 2010 GCAATGAG 2069
QY 2115 GGTGAG 2174
Db 2070 GGTGAG 2129
QY 2175 GATGAG 2234
Db 2130 GATGAG 2189
QY 2235 AACCTGAG 2294
Db 2190 AACCTGAG 2249
QY 2295 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2354
Db 2242 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2309
QY 2352 TCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2411
Db 2302 TGAATGAG 2361

QY 2412 TAAAAA 2434
 Db 2362 TTGCTACTTAAAAA 2384

RESULT 15
 DB3075
 Locus
 DEFINITION Lithospermum erythrorhizon mRNA for phenylalanine ammonia-lyase,
 complete cds.
 ACCESSION D83075
 VERSION D83075.1 GI:2911121
 KEYWORDS phenylalanine ammonia-lyase.
 SOURCE Lithospermum erythrorhizon dedifferentiated cells cell suspension
 cultures cell line:M8 cDNA to mRNA, clone:lib:expression library
 of lithospermum cells cultured in pigment production medium
 clone:LEPAL-1.

ORGANISM
 Lithospermum erythrorhizon

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; euasterids I incertae sedis; Boraginaceae;
 Lithospermum.

REFERENCE
 1 (sites)

Yazaki, K., Kataoka, M., Honda, G., Severina, K. and Heide, L.
 CDNA cloning and gene expression of phenylalanine ammonia-lyase in
 Lithospermum erythrorhizon
 Biosci. Biotechnol. Biochem. 61 (12), 1995-2003 (1997)

REFERENCE
 2 (bases 1 to 2441)

Yazaki, K.
 Direct Submission
 Submitted (16-Jan-1996) Kazufumi Yazaki, Graduate School of
 Agriculture, Kyoto University, Lab. Molecular & Cellular Biology,
 div. Applied Life Sciences, Kitashitakawa, Kyoto, Kyoto 606-01,
 Japan (E-mail: yazaki@kais.kyoto-u.ac.jp, Tel: 075-753-6384,
 Fax: 075-753-6398)

FEATURES
 source

Location/Qualifiers
 1..2441
 /organism="Lithospermum erythrorhizon"
 /db_xref="taxon:34254"
 /clone="LEPAL-1"
 /cell_line="M18"
 /tissue_type="cell suspension cultures"
 /clone_lib="expression library of Lithospermum cells
 cultured in pigment production medium"
 /dev_stage="differentiated cells"
 /note="Site. et zucc."
 161..2293

CDS

/codon_start=1
 /product="phenylalanine ammonia-lyase"
 /protein_id="AAA24928.1"
 /db_xref="GI:2911122"

/translation="MFTIVENGCKTMEFCMKDPLNNEMASFSKMGSHLDEKNNVAE
 FRKPVOLAGKITIGQVAAIARDGVTELEAAEGYKASDWMDSMNGTDSY
 GVTGEGATSHRTKGGALQKELIRFLNMGIPFNGTETSHLPHSATPAMVPIPT
 LLOGSGIRFELFAITKFLNTNITPCLPRGTTASGDVPLSTAGLTGSPNKA
 VCGPCEKINAEARLACISGTFEIDPKESALVNTGASGMSAVLYEALILAVL
 SEVISATFAEYWNKPEPTDHLRKLHHPGQIPAAIMHIIIDGSOYAAOKLHEM
 NFOGPIKODRYALRTSPQWIGPOLLEVIISATKMIEREINSVNPIIDVSRNALHG
 GAEIMASYCESELOFANPTVNHQSAEONHODVNSGLISKRTSPAVIILKMSF
 FLVLFQAVDRHTEENYRLAVKNTVSQVKKRITLTGYNELHSPSEEDILRYDR
 EYVAYADPCITTYPIIMOKIRELTVCHALDNGNEDVNTSIFPKAIAFEREIKAIL
 PKVHNARASVENCIPATSNRIEGRSTPYLKVYRELGTELLTGKRVSPGHEIDKY
 FTAKCEKLVADPLACLFAMNGAPLPIC"

polyA_signal
 BASE COUNT 663 a 526 c 567 g 685 t
 ORIGIN

Query Match 50.3%, Score 1227.4, DB 8, length 2441;
 Best local Similarity 72.9%, Prod. No. 2.8e-283;
 Matches 1595; Conservative 0; Mismatches 586; Indels 6; Gaps 1;

QY 135 AGCTAATGAGATGCTTATATAGTGTGATCATAGATCATTAATGATGAGAGTTCAG 194
 Db 180 ATGGAATATGAAAAAACTATGAGATTTCATGAGAAAGATCATTAATGAGAGATG 239
 QY 195 CGAGCGCTTGACCGGAGAGTCACTTGTATGAGTGAAGAAATGTTGGCAGCTTCAGAA 254
 Db 240 CTGAGTCATGAAAGGAGACACCTTGATGAAAGTAAAGAAATGATGAGTTCAGAA 299
 QY 255 ACCCGGTGTGAGAGTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 314
 Db 300 AACCGGTGTGAGAGTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 359
 QY 315 CTGCTATATACAGTATACACACCGTGAAGTGTGATGAGAGAGAGAGAGAGAGAGAG 374
 Db 414 AGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 473
 QY 435 CCACCGGCTTGCGGCGCACCTTCACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 494
 Db 474 CCACCGGCTTGCGGCGCACCTTCACCTTACATAGAGAGAGAGAGAGAGAGAGAGAG 533
 QY 495 AGCTATATGATTTTGAAGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 554
 Db 534 AACTTATATGATTTTGAAGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 593
 QY 555 TTGCATTTAGAGTTCAG 614
 Db 594 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 653
 QY 615 ACACCGGCTTGCGGCGCACCTTCACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 674
 Db 654 ATTCAAGCATGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 713
 QY 675 CCCCTGTTTACCCCTCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 734
 Db 714 CTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 773
 QY 735 ACATGCGGCGCTTGAAG 794
 Db 774 ACATGCGGCGCTTGAAG 833
 QY 795 TCCCTAATGCGGAG 854
 Db 834 ACATGATGCGGAG 893
 QY 855 AGCGGAAG 914
 Db 894 AGCTTGAAG 953
 QY 915 TCGCTTATTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCT 974
 Db 954 TGGTTTATTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCT 1013
 QY 975 CTGAGGTATGCAAG 1034
 Db 1014 CTGAGGTATGCAAG 1073
 QY 1035 ACCCGGTGATGAG 1094
 Db 1074 ATCCAG 1133
 QY 1095 TCAAGCGGCGGCAAG 1154
 Db 1134 TTAAGGCTGCTAG 1193
 QY 1155 ATGCTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1214
 Db 1194 ATGCTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1253
 QY 1215 CCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1274

.

.

.

.

.

.

Genforce version 5.1.4.F5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 28, 2003, 21:47:27 : Search time 3113 seconds

(without alignments)
12712.754 Million cell updates/sec

Title: us-09-964-992a-3

Perfect score: 2442

Sequence: 1 gqcaalcgcalacaccc aaaaaaaaaaaaaaaaaa 2442

Scoring table: IDENTITY_NUC Gapop 10.0 , Gapext 1 0

Searched: 16154066 seqs, 800774376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 98

Maximum Match 100%

Listing first 45 summaries

Database :

1: em_estha:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estrov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: qb_est1:*
10: qb_est2:*
11: qb_hic:*
12: qb_est3:*
13: qb_est4:*
14: qb_est5:*
15: em_estfun:*
16: em_estom:*
17: qb_gss:*
18: em_gss_hum:*
19: em_gss_luv:*
20: em_gss_fth:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_alter:*
26: em_gss_pro:*
27: em_gss_tod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	942.4	38.6	2598	11	AY103647 Zea mays
2	857.6	35.1	2299	11	AY106831 Zea mays
3	759	31.1	759	14	HO850732 Zea mays
4	758	31.0	1864	11	AY104679 Zea mays
5	715	29.3	741	14	BO941054 Zea mays
6	708.2	29.0	770	14	HO999214 Zea mays

7	694.4	28.4	710	14	BO869202
8	682.6	28.0	712	14	BO869225
9	661.4	27.1	679	14	HO998011
10	657	26.9	661	14	HO869819
11	656.8	26.9	679	14	HO007019
12	623	25.5	655	14	BO001313
13	564.2	23.1	674	14	BO865594
14	526.8	21.6	711	14	BO007176
15	523.8	21.4	696	14	HO024517
16	518.8	21.2	634	14	HO998720
17	518.8	21.2	802	14	HO645264
18	514.6	21.1	697	14	BO028965
19	513.6	21.0	707	14	BO028661
20	508.2	20.8	807	14	BO929414
21	507.2	20.8	707	14	BO028321
22	506.2	20.7	715	14	BO028594
23	500.6	20.5	692	14	HO029550
24	498.8	20.4	779	12	BO351058
25	496.8	20.3	713	14	BO027265
26	496.8	20.3	792	12	BO645267
27	494.6	20.3	799	12	HO887997
28	494	20.2	806	13	BO406502
29	491.4	20.1	687	14	HO028544
30	490.6	20.1	676	14	BO028141
31	490.4	20.1	508	14	BO987939
32	486.8	19.9	699	14	BO029186
33	476.8	19.5	819	12	HO003440
34	472.2	19.3	767	14	BO045673
35	471.6	19.3	479	14	HO998697
36	471.4	19.3	734	13	BO933291
37	471.2	19.3	726	10	AM218834
38	471	19.3	624	14	BO968511
39	469	19.2	631	14	HO003880
40	468.2	19.2	626	14	HO004978
41	467.4	19.1	613	14	HO998679
42	467.4	19.1	631	12	BO925579
43	465.6	19.1	794	12	BO647641
44	465.2	19.0	639	14	BO027438
45	464.6	19.0	744	12	HO597014

ALIGNMENTS

BO869202 CDS:112..Y
BO869225 CDS:1010..Y
HO998011 CDS:1805..Y
HO869819 CDS:721..Y
HO007019 CDS:4101..Y
BO001313 CDS:2781..Y
BO865594 CDS:522..Y
BO007176 CDS:1311..Y
HO024517 CDS:607..Y
HO998720 CDS:2822..Y
HO645264 CDS:506884..Y
BO028965 CDS:112..Y
BO028661 CDS:105..Y
BO929414 CDS:4933..Y
BO028321 CDS:1201..Y
BO028594 CDS:10..Y
HO029550 CDS:1214..Y
BO351058 CDS:609..Ma
BO027265 CDS:302..Y
BO645267 CDS:506884..Y
HO887997 CDS:13848..Y
BO406502 CDS:580829..Y
HO028544 CDS:15012..Y
BO028141 CDS:7821..Y
BO987939 CDS:1309..Y
BO029186 CDS:9111..Y
HO003440 CDS:31948..Y
HO045673 CDS:594791..Y
HO998697 CDS:1011..Y
BO933291 CDS:53140..Y
AM218834 CDS:301314..Y
BO968511 CDS:34614..Y
HO003880 CDS:36113..Y
HO004978 CDS:6601..Y
HO998679 CDS:2706..Y
BO925579 CDS:138..StA
BO647641 CDS:509260..Y
BO027438 CDS:5810..Y
HO597014 CDS:95692..Y

RESULT 1
LOCUS AY103647
DEFINITION Zea mays PC0142084 mRNA sequence.
ACCESSION AY103647
VERSION AY103647.1 GI:2126725
KEYWORDS
SOURCE
ORGANISM Zea mays.
Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta, Spermatophyta, Magnoliophyta, Liliopsida, Poales, Poaceae, PACC
Clade: Panicoideae; Andropogoneae; Zea.
REFERENCE
AUTHORS Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S., Arthur, L.W., Hainey, M., Morgante, M. and Tingey, S.V.
TITLE Maize Mapping Project/Dupont Consensus Sequences for Design of Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE
AUTHORS Coe, E.C.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA
FEATURES
Source
1..2598
/organism="Zea mays"
/db_xref="MaizeDB:638378"
/db_xref="taxon:4577"
/clone="PC0142084"

/clone_11b="Maize Mapping Project/DuPont Cornsensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed Dupont contigs; this resource was
assembled by Dupont as part of a collaboration for the
overgo addressing of BACS in conjunction with the Maize
Mapping Project"

BASE COUNT 490 a 907 c 795 g 406 t
ORIGIN
Query Match 38.6%; Score 942.4; DB 11; Length 2598;
Best Local Similarity 66.6%; Pred. No. 9.2e-115;
Matches 1383; Conservative 0; Mismatches 686; Indels 9; Gaps 2;

170 GATCACTTAAGTGGGAGTTGACAGGCGGTTGACCGGACGACGCTGATGAGGTG 229
140 GACCCGCTAAGTGGGAGGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGG 199
230 AAGAGATGTTGCGGAGTTGAGAAAGCGGTTGTTGAGTGGAGGAGGAGGAGG 289
200 AAGCGGATGTTGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 259
290 GTTTCAGAGTGGGCGGAGTGGGAGGCTGATGAGGAGGAGGAGGAGGAGGAG 349
260 GTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 319
350 TCGGAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 409
320 GAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 379
410 AAGAGATGTTGAGTATGTTGAGTATGTTGAGTATGTTGAGTATGTTGAGT 469
380 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 439
470 AAGAGATGTTGAGTATGTTGAGTATGTTGAGTATGTTGAGTATGTTGAGT 529
440 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 499
530 AATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 589
500 ACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 556
590 AGATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 649
557 CGATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 616
650 ACCAGATGTTGAGTATGTTGAGTATGTTGAGTATGTTGAGTATGTTGAGT 709
617 ACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 676
710 TCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 769
677 TCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 736
770 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 829
737 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 796
830 GTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 889
797 ATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 856
890 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 949
857 TCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 916
950 TCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1009
917 TCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 976
1010 CACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1069

977 CACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1036
1070 TATATTTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1129
1037 CACATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1096
1130 TATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1189
1097 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1156
1190 CAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1249
1157 CAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1216
1250 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1309
1217 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1276
1310 ACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1369
1277 ACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1336
1370 ATGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1429
1337 ATGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1396
1430 CTCTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1489
1397 CTCTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1456
1490 GCTTCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1549
1457 GCTTCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1516
1550 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1609
1517 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1576
1610 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1669
1577 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1636
1670 ATGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1729
1637 GTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1696
1730 GTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1789
1697 GTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1756
1790 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1849
1757 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1816
1850 TATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1909
1817 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1876
1910 AACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1969
1877 AACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1930
1970 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2029
1931 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1990
2030 AATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2089
1991 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2050
2090 TTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2149
2051 TTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2210

University of California at Davis (UCD)
 Asmundson Hall, mcd, Davis, CA 95615, USA
 Tel: 1 (530) 742-1742
 Fax: 1 (530) 752-9659

Email: akozik@ucdavis.edu [michelmore@ucdavis.edu]
 belongs to contig q9_CA.Contig196760, see <http://cgpdb.ucdavis.edu/>
 for details.
 Plate: q399, row: N, column: 08.

FEATURES

Source

Location/Qualifiers

1..770
 /organism="Lactuca sativa"
 /cultivar="L.serriola"
 /db_xref="taxon:4236"
 /clone="OG99N08"
 /clone_lib="OG_FPCBU lettuce serriola"

/note="Vector: pBRCDNA51AB; The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformation made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at <http://cgpdb.ucdavis.edu/>
 TAG_SEQ=Not found"

BASE COUNT 188 a 211 c 200 g 171 t

ORIGIN

Query Match

29.0%; Score 708.2; DB 14, Length 770,
 Best Local Similarity 98.8%, Pred. No. 9, 9e 84,

Matches 724; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 390 GCGTATGACGACATGAATAAGCAAGTATGATGCTGACACCGGCTTCGGG 449
 Db 39 GGGTATGACGACATGAATAAGCAAGTATGATGCTGACACCGGCTTCGGG 98
 QY 450 GCGCTTCTTATGACGACATGAATAAGCAAGTATGATGCTGACACCGGCTTCGGG 509
 Db 99 GCGCTTCTTATGACGACATGAATAAGCAAGTATGATGCTGACACCGGCTTCGGG 158
 QY 510 TGAAGCTTCTTATGACGACATGAATAAGCAAGTATGATGCTGACACCGGCTTCGGG 569
 Db 159 TGAAGCTTCTTATGACGACATGAATAAGCAAGTATGATGCTGACACCGGCTTCGGG 218
 QY 570 CCAAGCGCGGATGACGACATGAATAAGCAAGTATGATGCTGACACCGGCTTCGGG 629
 Db 219 CCAAGCGCGGATGACGACATGAATAAGCAAGTATGATGCTGACACCGGCTTCGGG 278
 QY 630 TCAAGCTTCTTATGACGACATGAATAAGCAAGTATGATGCTGACACCGGCTTCGGG 689
 Db 279 TCAAGCTTCTTATGACGACATGAATAAGCAAGTATGATGCTGACACCGGCTTCGGG 338
 QY 690 TCAAGCTTCTTATGACGACATGAATAAGCAAGTATGATGCTGACACCGGCTTCGGG 749
 Db 339 TCAAGCTTCTTATGACGACATGAATAAGCAAGTATGATGCTGACACCGGCTTCGGG 398
 QY 750 TCAAGCTTCTTATGACGACATGAATAAGCAAGTATGATGCTGACACCGGCTTCGGG 809
 Db 399 TCAAGCTTCTTATGACGACATGAATAAGCAAGTATGATGCTGACACCGGCTTCGGG 458
 QY 810 TCAAGCTTCTTATGACGACATGAATAAGCAAGTATGATGCTGACACCGGCTTCGGG 869
 Db 459 TCAAGCTTCTTATGACGACATGAATAAGCAAGTATGATGCTGACACCGGCTTCGGG 518
 QY 870 TCAAGCTTCTTATGACGACATGAATAAGCAAGTATGATGCTGACACCGGCTTCGGG 929
 Db 519 TCAAGCTTCTTATGACGACATGAATAAGCAAGTATGATGCTGACACCGGCTTCGGG 578
 QY 930 TCAAGCTTCTTATGACGACATGAATAAGCAAGTATGATGCTGACACCGGCTTCGGG 989
 Db 579 TCAAGCTTCTTATGACGACATGAATAAGCAAGTATGATGCTGACACCGGCTTCGGG 638

QY 990 GGAAGCTTCTTATGACGACATGAATAAGCAAGTATGATGCTGACACCGGCTTCGGG 1049
 Db 639 GGAAGCTTCTTATGACGACATGAATAAGCAAGTATGATGCTGACACCGGCTTCGGG 698
 QY 1050 GGAAGCTTCTTATGACGACATGAATAAGCAAGTATGATGCTGACACCGGCTTCGGG 1109
 Db 699 GGAAGCTTCTTATGACGACATGAATAAGCAAGTATGATGCTGACACCGGCTTCGGG 757
 QY 1110 AGCTCCACGCAAT 1122
 Db 758 AGCTCCACGCAAT 770

RESULT 7

H0869202

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 710)
 Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Riesberg, L.,
 Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, D., Ellison,
 P., Kolkman, J., Slabaugh, M.S., Livingston, K., Zhou, Y., Lal, Z.,
 Church, S., Jackson, L. and Bradford, K.
 lettuce and sunflower ESTs from the Composite Genome Project
<http://comgenomics.ucdavis.edu/>
 unpublished (2002)
 Contact: Alexander Kozik [R.W.Michelmore]
 Department of Vegetable Crops, R.W.Michelmore Lab
 University of California at Davis (UCD)
 Asmundson Hall, UCD, Davis, CA 95616, USA
 Tel: 1-(530)-742-1742
 Fax: 1-(530)-752-9659
 Email: akozik@ucdavis.edu [michelmore@ucdavis.edu]
 belongs to contig q9_CA.Contig196760, see <http://cgpdb.ucdavis.edu/>
 for details.
 Plate: q655, row: H, column: 12.

FEATURES

Source

Location/Qualifiers

1..710
 /organism="Lactuca sativa"
 /cultivar="Salinas"
 /db_xref="taxon:4236"
 /clone="OGD5H12"
 /clone_lib="OG_AHCD1 lettuce salinas"

/note="Vector: pBRCDNA51AB; The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformation made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at <http://cgpdb.ucdavis.edu/>
 TAG_LIB=OG_AHCD1 lettuce salinas
 TAG_TISSUE=Chemical induction
 TAG_SMO=HCTGCCCGG"

BASE COUNT

191 a 174 c 171 g 174 t

ORIGIN

Query Match

28.4%; Score 694.4; DB 14, Length 710,
 Best Local Similarity 99.7%, Pred. No. 6, 7e 82,

Matches 706; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 830 GTTAACTTCTTATGACGACATGAATAAGCAAGTATGATGCTGACACCGGCTTCGGG 889

Email: akozik@atgc.org [michelmore@vegmall.ucdavis.edu]
 belongs to contig OG_CA.Contig6760, see http://cgpdb.ucdavis.edu/
 for details.

FEATURES
 source

Location/Qualifiers

1..661
 /organism="Lactuca sativa"
 /cultivar="Salinas"
 /db_xref="taxon:4236"
 /clone="OGSD7E21"
 /clone_lib="OG_ABCDI lettuce salinas"
 /lab_host="E.coli"
 /note="Vector: pBRCDNA51AB: The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpdb.ucdavis.edu/
 TAG_LIB-OG_ABCDI lettuce salinas
 TAG_TISSUE="chemical induction
 TAG_SBO="GTACCCGGC"

BASE COUNT 160 a 185 c 178 g 138 t

Query Match 26.9%; Score 657; DB 14; Length 661;

Best Local Similarity 100.0%; Pred. No. 5,4e-77;

Matches 657; Conservative 0; Mismatches 0; Indels 0; Caps 0;

OG 268 CCTCGGAGGAGAGACGCTTACAGTTCTTCAGGTGGCGGGGATCGACCTGCTAATGACAG 327
 |||||||
 Db 1 CGCGGAGGAGAGAGCTTACAGTTCTTCAGGTGGCGGGGATCGACCTGCTAATGACAG 60
 |||||||
 OG 328 TGACACCGCTGAAGGTGGAGCTGTGCGAAGCCGCGGCTGAGCTTAAGCGGAGTAGTGA 387
 |||||||
 Db 61 TGACACCGCTGAAGGTGGAGCTGTGCGAAGCCGCGGCTGAGCTTAAGCGGAGTAGTGA 120
 |||||||
 OG 388 TTGGGTTATGAGACATGATTAAGAACTGATAGTATGATGATGATGATGATGATGATG 447
 |||||||
 Db 121 TTGGGTTATGAGACATGATTAAGAACTGATAGTATGATGATGATGATGATGATGATG 180
 |||||||
 OG 448 CGCGACCTTCACCGGAGAGCTTAAGAACTGATAGTATGATGATGATGATGATGATGATG 507
 |||||||
 Db 181 CGCGACCTTCACCGGAGAGCTTAAGAACTGATAGTATGATGATGATGATGATGATGATG 240
 |||||||
 OG 508 TTTCACCGCTTCACCGGAGAGCTTAAGAACTGATAGTATGATGATGATGATGATGATG 567
 |||||||
 Db 241 TTTCACCGCTTCACCGGAGAGCTTAAGAACTGATAGTATGATGATGATGATGATGATG 300
 |||||||
 OG 568 CATTCAGATCTTGAAGACCATGACCAAGTTCCTTAACAAACAATCAACCCCTTTTACG 627
 |||||||
 Db 301 CATTCAGATCTTGAAGACCATGACCAAGTTCCTTAACAAACAATCAACCCCTTTTACG 360
 |||||||
 OG 628 ATTTCAGATCTTGAAGACCATGACCAAGTTCCTTAACAAACAATCAACCCCTTTTACG 687
 |||||||
 Db 361 ATTTCAGATCTTGAAGACCATGACCAAGTTCCTTAACAAACAATCAACCCCTTTTACG 420
 |||||||
 OG 688 CCTTCAGATCTTGAAGACCATGACCAAGTTCCTTAACAAACAATCAACCCCTTTTACG 747
 |||||||
 Db 421 CCTTCAGATCTTGAAGACCATGACCAAGTTCCTTAACAAACAATCAACCCCTTTTACG 480
 |||||||
 OG 748 CTTTACCGCGCGCGCTTAAGAGCTTGAAGAGCTTGAAGAGCTTGAAGAGCTTGAAGAG 807
 |||||||
 Db 481 CTTTACCGCGCGCGCTTAAGAGCTTGAAGAGCTTGAAGAGCTTGAAGAGCTTGAAGAG 540
 |||||||
 OG 808 AAAGGCTTCGCTGACCGCGAGTGAAGGTGGGTTCTTCAGAGTACAGCGGAAAGAGG 867
 |||||||
 Db 541 AAAGGCTTCGCTGACCGCGAGTGAAGGTGGGTTCTTCAGAGTACAGCGGAAAGAGG 600
 |||||||
 OG 868 GCTACGATTTTAAAGCGACCGCGGTGGGTTCCGCGATGCTTCATGCTTCTATT 924
 |||||||

Db 601 CCTACACATCTTGAAGAGCGACCGCGCGCGCGCGAGTGCCTTCAATGCTTATT 657

RESULT 11

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

ORGANISM

SOURCE

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

location/Qualifiers

1..679

/organism="Lactuca sativa"

/cultivar="L.serrifolia"

/db_xref="taxon:4236"

/clone="OGH4H01"

/clone_lib="OG_FFGHJ lettuce serrifolia"

/lab_host="E.coli"

/note="Vector: pBRCDNA51AB: The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpdb.ucdavis.edu/
 TAG_LIB-OG_FFGHJ lettuce serrifolia
 TAG_TISSUE="leaves dark grow
 TAG_SBO="GCTACGCGC"

BASE COUNT 190 a 145 c 186 g 158 t

Query Match 26.9%; Score 656.8; DB 14; Length 679;

Best Local Similarity 98.8%; Pred. No. 5.6e-77;

Matches 672; Conservative 0; Mismatches 7; Indels 1; Caps 1;

OG 1 GAGCAATCTGATCATATACCATTCATGACGATCAAGAGTGTATGATGATGATGATGATG 60
 |||||||
 Db 1 GAGCAATCTGATCATATACCATTCATGACGATCAAGAGTGTATGATGATGATGATGATG 60
 |||||||
 OG 61 ACACATTAAGATGTTCTTGTCTTTCATATAGTCTACAAATCTGTAAATAATAT 120
 |||||||
 Db 61 ACACATTAAGATGTTCTTGTCTTTCATATAGTCTACAAATCTGTAAATAATAT 120
 |||||||
 OG 121 GGAGAAAGGTAATCAAGTTAATGAGTGTATGATGATGATGATGATGATGATGATGATG 180
 |||||||
 Db 121 GGAGAAAGGTAATCAAGTTAATGAGTGTATGATGATGATGATGATGATGATGATGATG 180
 |||||||

VERSION	EST.					
KEYWORDS	H0865594.1 GI:2251059					
SOURCE	Lactuca salvia.					
ORGANISM	Lactuca salvia. Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons, core eudicots, Asteridae; euasterids II; Asterales; Asterales; Lactuceae; Lactucaceae; Lactuca.					
REFERENCE	1 (bases 1 to 674)					
AUTHORS	Kozik,A., Michelmore,R.W., Knapp,S., Malvinenko,M., Riesberg,L., Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,T., Ellison ,P., Kolman,J., Slabaugh,K.S., Livingston,K., Zhou,Y., Lai,Z., Church,S., Jackson,I., and Bradford,K. Lettuce and Sunflower ESTs from the Compositae Genome Project http://compgenome.ucdavis.edu/ Unpublished (2002) Contact: Alexander Kozik [P W Michelmore-] Department of Vegetable Crops, R.W.Michelmore lab University of California at Davis (UCD) Asmundson Hall,UCD, DAVIS, CA 95616, USA Tel.: 1-(530)-742-1742 Fax: 1-(530)-752-9659 Email: akozik@ucdavis.org [michelmore@vegemail.ucdavis.edu] belongs to contig QC_CA.Contig9517, see http://cgpbdb.ucdavis.edu/ for details. [plate: QCCS, row: F column: 22.]					
JOURNAL						
COMMENT						
TITLE						
FEATURES	Location/Qualifiers					
Source	1..674 /organism="Lactuca salvia" /cultivar="Salinas" /db_xref="taxon:4216" /clone="QCCSF22" /clone_lib="QG_ABCDI lettuce salinas" /lab_host="E.coli" /note="Vector: pBREDNStAB. The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpbdb.ucdavis.edu/ TAbLib-QC_ABCDI lettuce salinas TAG_SEQ-Chemical induction TAG_SEQ-TGtagCGGG"					
BASE COUNT	167 a	179 c	172 g	156 t		
ORIGIN						
Query Match	23.1%	Score 564.2	DB 14,	Length 674;		
Best Local Similarity	89.9%	Pred. No 6,-9e-65;				
Matches 605;	Conservative 0;	Mismatches 68;	Indels 0;	Gaps 0		
Db	371 GTTTAAGCGCATATGTATTGGCTTATGGAGACATGAATAAAGAACAATGATTATGCT	430				
Db	2 GTCAAGCGCATATGTATTGGCTTATGGAGACATGAATAAAGAACAATGATTATGCT	61				
OY	431 GTCACCAAGCGCTTGCGGCACACTCTTCATCGGAAACTTAACCAAGCGCGTCTTAAAG	490				
Db	62 GTTCACTACCGCGTTTCGGAAGCTACCCCTCACCGGAGAACCAAGAGTGTCCTTTCAG	121				
OY	491 AAGAGCTCATATTAGATTTTTGAACGCCGGAATTTTGGCAATGACGAAGCAAGCAAGC	550				
Db	122 AAGAGCGCATATTATCTTGAAGGAGAGAGAGCAAGCCTTGTTCAGATCAACACCCTCTCAA	181				
OY	551 ACACCTTCCACATTTACGCCACACAGCGCGCATATGTCAGATCAACACCCCTCTCCAG	610				
Db	182 ACACCTTCCACATTTACGCCACACAGCGCGCATATGTCAGATCAACACCCTCTCAA	241				
OY	611 GGTACTTCGGGCATTCGATTCGATCTTGGAAAGCATATCGAAGTCTTTAAACAAGC	670				
Db	242 GGTACTTCGGGCATTCGATTCGATCTTGGAAAGCATATCGAAGTCTTTAAACAAGC	301				

QY	671	ATACCCCTTGTTACCCCTCGGTGAACCAATACACCGCTCCCGGTGACCTGGTCCCATTA	730
Db	302	GTACACCTTTTTTCTCTCTTCTGTGAGAAATTAATTAATCTCGGTGATCTGTCACATTA	361
QY	741	TCATATATGAGGAGCTCTCTTAACAGGAGTGGCTATATCTGAATATCTTGGTAAAGAA	790
Lb	362	TCTTACATGCGCGCTCTCTCAACCGCGCTGCGACATGCAAGACCGCTTGGACCCACCGGA	421
QY	791	GAAGTCTCTCAATGCTGAAGAAAGGCTTTGGTGTGAGGTGAATTTGAAGGTGAGTGT	850
Db	422	GAAGTCTTAAATATGCTGAAGAAAGGCTTTGGTGTGAGGTGAATTTGAAGGTGAGTGT	481
QY	851	TTACAGCGCAAGAGAGGCTAGACATCTGTAAACCGCACCGCGCGCTGCGGAGATGGCG	910
Db	482	TTACAGCGCAAGAGAGGCTAGACATCTGTAAACCGCACCGCGCGCTGCGGAGATGGCG	541
QY	911	TCATAGGTTCAATTTATATATGCTATGACTCTGGTGTGTGTGGAAGAGTATATGAGGATC	970
Db	542	TGCATGCTGCTATTTTGATGCTATATGTCGTGCAATGCTTGCGGAAGTGTATCGCGCATC	601
QY	971	TTGCGTAGGTATATGAAGGAAGGTGGAATTATCTGATATCTGGAATATGAATTTGAAG	1030
Db	602	TTGCGTAGGTATATGAAGGAAGGTGGAATTATCTGATATCTGGAATATGAATTTGAAG	661
QY	1031	CATCACCCCTGCTC 1043	
Db	662	CATCACCCCGGCTC 674	

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
B0007176	OCH1K13.yr.abl_QG-EFGH lettuce serriola Lactuca sativa cDNA clone	B0007176	B0007176.1	GJ:2441571	EST.	Lactuca sativa.
						Lactuca sativa
						Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
						Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
						Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae;
						Lactuca.
						1 (bases 1 to 711)
REFERENCE	Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L., Lin,H., van Damme,M., Lavelle,D., Chevallier,P., Ziegler,J., Ellison P., Kolkmann,T., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z , Church,S., Jackson,L. and Bradford K. lettuce and sunflower ESTs from the Composite Genome Project http://comgenomics.ucdavis.edu/ Unpublished (2002)					
JOURNAL	Contact: Alexander Kozik [P W Michelmore] Department of Vegetable Crops, R.W.Michelmore Lab University of California at Davis (UCD) Asmundson Hall, UCD, Davis, CA 95616, USA Tel.: 1-(530)-742-1742 Fax: 1-(530)-752-9659 Email: akozika@ucdavis.org [michelmore@veggemail.ucdavis.edu] Belongs to conlig QG-CA.Conlig96267, see http://cgpdb.ucdavis.edu/ for details.					
COMMENT	Plate: QGH13 row: K column: 13. Location/Qualifiers 1..711 /organism="Lactuca sativa" /cultivar="L.serriola" /db_xref="taxon:4236" /clone_id="QG13K13" /clone_lib="QG-EFGH lettuce serriola" /lab_host="E.coli"					
FEATURES	/note=Vector: pBRCMDNA51AB: The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each					
SOURCE						

HASH: COUNT	ORIGIN
169	a
192	c
194	g
152	t
4	others

best local similarity 84.26; Pseq. NO. 38-60;
Matches 591; Conservative 0; Mismatches 111; Indels 0; Gaps 0.

pb 130 CACTGTGAGCTGGAGCTGTCCAGACTTCCTATTCCTTGGATCAAGCGACTTACATCATTTG 189

Db 190 GGTGATGGAGAGCATATGAGGAGAGGATTCATATTTATGAGATTAATAGAGTTTGAGAG 249

Db 250 CACCTCTGACCCGAGGACCAAGCGGTGCTCTTCAAGAAGACTCAT¹TAGATTCCT 309

11b 310 AACCGCGGAATTNTTGGCAACGGAACGGATCAAGCCACACCCCTTGGCAATTACGCAAC 369

D0 370 C A6A8N9G6ALIGP GELAMZAIICAN A8CCEIPEE M88BILIN T8888ATF06GITT 429

691 CCGTGAATATCACTGAGTTGTGATTACATAATGGAGGCTTCCTC

751 AACCGGCGGCGCACTTCGAAGCCGTTGGCCCGACCGAGAGTTCCTCAATGCCGAAAA 810

QY 811 GGCCCTTCGCTCCAGCCGAGCTTCAAGCTTCGGCTTCCTCGACTTACACCCGAAGAAGACGGCT 870

871 AAGATTGTTAACTGGAGACGGAGGTGAGAGCTGGAGTATATTTC 912

Figure 15

Accession	Gene	Accession	Gene	Accession	Gene
Q9H4E07	yg.abl	Q9_HFHJH	sunflower	RHA280	Helianthus annuus cDNA
Q9H4E07	clone	Q9H4E07	mRNA structure		

KEYWORDS EST.
SOURCE common sunflower.

ORGANISM *Helianthus annuus*
 FUKARYOTA: Viridiplantae, Streptophyta; Embryophyta; Tracheophyta;

Liu, H., van Damme, M., Lavelle, B., Chevalier, P., Ziegler, J., Ellison, P., Kolman, J., Slabaugh, M. S., Livingston, K., Zhou, Y., Lai, Z.,

JOURNAL Unpublished (2002)
COMMENT Contact: Alexander Kozik (R.W.Michelmores)

Asmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-0650

email: dkozlik@atgc.org [michu.more@vegmil.ucdavis.edu]
 belongs to config QH_CA-Config3142, see <http://cgrpdb.ucdavis.edu>

source
I. 696
/organism-"Helianthus annuus"

```

/closure="QH4e0/"
/closure_lib="QH_EFGH synflower PHA280"
/closure_lib="QH_EFGH synflower PHA280"

```

from 11 different sources of RNA from a single genome. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each

transformations made with four size classes to minimize size bias. Details of each source of RNA and library

TAG_TISSUE-chemical induction
TAG_SEQ-TGTAGCCGGG"

Query Match	21.48;	Score 523.8;	DB 14;	Length 696;
Best Local Similarity	84.68;	Prefd No 1	3e-59;	

521 ATATTGGCATGGACGGAAACCAAGCCACACACTTCCACATTCCAGCCACGAGCGGCC 5,

61 A C C T T G C T A G A A T C A A C A C C C T C C T C C A G G A A T A T T C G C C A T A A G A A T T T C G A A A T C T T G G 1

Db 121 GAAGCATTACTAAATGCTTAAACACATATGCTTATCTGAGAGAG 1

701 ATGACCGGGCTTGCAGTTGTCCCAATTAICAAACATCGGCGGCGCTATTAAAGCGGCCG

24) GCGAACTGCAAAAGTACTTGCCGTGGAGGAGGATGAAGTTCTTCACATGAGTAATGACCATTTGCCCGC 30

301 CAGCTGGCTATGAAAGTGGATTTCGACGTTAACTTAAACAAAGAGTCGGCTGT 302

881 AACGGCACCCCGCTGGGGTCCGGGATGGCTTCATGCTGCTTAATGTAATT 9

```
Db 361 AATGATACCGGCGGTGGGCTGCGGATGAGCTGCTATGTTGTTTGAAGGCAATGTACTT 420
QY 941 GCGTGTGTTGCGGAAGTGTATCGGCGCATCTTCGCTGAGGTTATGCAAGGCAAGCCGAG 1000
Db 421 GCATTGTGTGGAGTGTGTGCGCATTTTTCGCGAATATGCAAGGCAAGGCGGAG 480
QY 1001 TTACCGATCACTTGACACACAAATTTGAAGCATCACTGCTGCTTAAATGAGAGGAGTGG 1060
Db 481 TTACCGACACATTTGACCCACAACTTAAGCACCAAGCGCGCAAAATAGAGCGCGCAGCC 540
QY 1061 ATCATGAGATATTTTGGAGCGAAGGATTACGTCAGAGCGCGCAAAAGTCCAGAA 1120
Db 541 ATCATGAGATATTTTGGAGCGAAGGATTACGTCAGAGCGCGCAAAAGTCCAGAG 600
QY 1121 ATGGAACCGTTACAGAAACCAAAACAAAGTCATGCTGCTGCTGCTGCTGCTGCTGCT 1180
Db 601 ATGGAACCGTTACAGAAACCAAAACAAAGTCATGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 1181 CTGGACCTCAAAATGGAAGTATCCGATCATCAAC 1215
Db 661 CTGGACCAAAATGGAAGTATCCGTCGCCGAC 695
```

Search completed: March 29, 2003, 00:34:00
Job time : 3134 secs

Db 61 ACACAAATAGATGTTCTTCTTCTTGTATAGTCTACAACTGTAATATATAT 120
QY 121 GCGAACCGTAATCAGCTTAATGAGTCTTATAGTGTGTCATCAAGATCCATTAA 180
Db 121 GCGAACCGTAATCAGCTTAATGAGTCTTATAGTGTGTCATCAAGATCCATTAA 180
QY 181 CTGGGGAGTTGCGAGGAGGCTTTCAGGAGAGTCACTTCATGAGGTGAAGATGCT 240
Db 181 CTGGGGAGTTGCGAGGAGGCTTTCAGGAGAGTCACTTCATGAGGTGAAGATGCT 240
QY 241 TCGGAGATTAGAAACCGGCTGTGAAGTCCGAGAGACAGCTTTACAGTTTTCAGCT 300
Db 241 TCGGAGATTAGAAACCGGCTGTGAAGTCCGAGAGACAGCTTTACAGTTTTCAGCT 300
QY 301 GCGGCGGATCGCAGCTGTCTAATGACAGTCAACCGTGAAGTGGAGCTGCGAAGCCGC 360
Db 301 GCGGCGGATCGCAGCTGTCTAATGACAGTCAACCGTGAAGTGGAGCTGCGAAGCCGC 360
QY 361 GAAGGCTGAGATTAAAGGAGTAGTATGAGTTATGAGAGCATGAAATAAAGAACTGA 420
Db 361 GAGGCGCTGAGATTAAAGGAGTAGTATGAGTTATGAGAGCATGAAATAAAGAACTGA 420
QY 421 TAGTATAGTGTGACAGCGGCTTCCGCGGCGACCTCTCCGAGAGACCTAACCAAGCGCG 480
Db 421 TAGTATAGTGTGACAGCGGCTTCCGCGGCGACCTCTCCGAGAGACCTAACCAAGCGCG 480
QY 481 TCGCTTACAGAAAGAGCTCAATAGATTGTTTGAACCGCGGAATATTCGGCAATGCAACGGA 540
Db 481 TCGCTTACAGAAAGAGCTCAATAGATTGTTTGAACCGCGGAATATTCGGCAATGCAACGGA 540
QY 541 AACAAAGCAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 600
Db 541 AACAAAGCAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 600
QY 601 CCTCTCTCAGAGTTTACCTCCGCGGCGGATTCGAGATCTTGAAGCCATCACCAGTTTCT 660
Db 601 CCTCTCTCAGAGTTTACCTCCGCGGCGGATTCGAGATCTTGAAGCCATCACCAGTTTCT 660
QY 661 TAAACAAACATCAACCCCTTTTACCCCTCCGTCGAAACCATCAGCCCTCCGCTGACT 720
Db 661 TAAACAAACATCAACCCCTTTTACCCCTCCGTCGAAACCATCAGCCCTCCGCTGACT 720
QY 721 TGCCCTATTATCATATCATATCATATCATATCATATCATATCATATCATATCATATCATAT 780
Db 721 TGCCCTATTATCATATCATATCATATCATATCATATCATATCATATCATATCATATCATAT 780
QY 781 CCCCACCGGAGAGTCTCTAATCCGAAAGGCTTCCGTCGAGCGGAGTTGAAAGTTCG 840
Db 781 CCCCACCGGAGAGTCTCTAATCCGAAAGGCTTCCGTCGAGCGGAGTTGAAAGTTCG 840
QY 841 GTTCTTCTGATATACAGCCCAAGAAAGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 900
Db 841 GTTCTTCTGATATACAGCCCAAGAAAGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 900
QY 901 CCGGAGTGTCTCTCATGTTTATTTATGTTATGTTATGTTATGTTATGTTATGTTATGTTAT 960
Db 901 CCGGAGTGTCTCTCATGTTTATTTATGTTATGTTATGTTATGTTATGTTATGTTATGTTAT 960
QY 961 ATCGCGGATCTTCTGAGGTTATGCAAGGAAAGGCTTCTTCTTCTTCTTCTTCTTCTTCT 1020
Db 961 ATCGCGGATCTTCTGAGGTTATGCAAGGAAAGGCTTCTTCTTCTTCTTCTTCTTCTTCT 1020
QY 1021 CAATTTGAACATCACCCTCTCTCTCAATTCGAGCGCGCGGATCATGAGTATATTTTGA 1080
Db 1021 CAATTTGAACATCACCCTCTCTCTCAATTCGAGCGCGCGGATCATGAGTATATTTTGA 1080
QY 1081 GCGAAGGATTAAGTAAAGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1140
Db 1081 GCGAAGGATTAAGTAAAGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1140
QY 1141 AAAACAAGATGTTATGCT 1200
Db 1141 AAAACAAGATGTTATGCT 1200

QY 1201 AATCGATCATCAACCAATGATGAGAGGCAATCAATTCCGCTCAATCAATCCATT 1260
Db 1201 AATCGATCATCAACCAATGATGAGAGGCAATCAATTCCGCTCAATCAATCCATT 1260
QY 1261 GATGAGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1320
Db 1261 GATGAGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1320
QY 1321 AGTTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCT 1380
Db 1321 AGTTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCT 1380
QY 1381 ATTTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCT 1440
Db 1381 ATTTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCT 1440
QY 1441 ACCTAACCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCT 1500
Db 1441 ACCTAACCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCT 1500
QY 1501 TTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCT 1560
Db 1501 TTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCT 1560
QY 1561 CAATCAAGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCT 1620
Db 1561 CAATCAAGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCT 1620
QY 1621 CATCTTAAACATCATGCTGCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCT 1680
Db 1621 CATCTTAAACATCATGCTGCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCT 1680
QY 1681 CCATTTGAGAGAGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1740
Db 1681 CCATTTGAGAGAGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1740
QY 1741 GGTCTTCAACATGAGGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCT 1800
Db 1741 GGTCTTCAACATGAGGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCT 1800
QY 1801 CCTCTGAGTTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860
Db 1801 CCTCTGAGTTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860
QY 1861 ATACCCATTAAATGAGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCT 1920
Db 1861 ATACCCATTAAATGAGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCT 1920
QY 1921 AAGGAGAAAGAACTCAACCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCT 1980
Db 1921 AAGGAGAAAGAACTCAACCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCT 1980
QY 1981 GAAAGTCTGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2040
Db 1981 GAAAGTCTGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2040
QY 2041 GTTCTTCTGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2100
Db 2041 GTTCTTCTGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2100
QY 2101 GAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2160
Db 2101 GAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2160
QY 2161 CAGGCTGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2220
Db 2161 CAGGCTGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2220
QY 2221 AGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2280
Db 2221 AGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2280


```

: SEQ ID NO 29
: LENGTH: 2163
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: Preferred
: OTHER INFORMATION: Theoretical sequence based in part on SEQ ID NO:20
: NAME/KEY: CDS
: LOCATION: (1)..(2163)
: OTHER INFORMATION:
US-09-939-408A-29

```

```

Query Match      10.9% Score 265.6; DB 10; Length 2163;
Best Local Similarity 51.6%; Pred. No. 1.1e-54;
Matches 647; Conservative 33; Mismatches 562; Indels 12; Gaps 4;

```

```

QY 549 ACACACTTCACATTCAGCCACAGAGCCGCCATTCCTGCAATACACACCTCCCTCC 608
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 491 ACTGCTTCCGCTCGAGATGTTCTGCGGATATGATATCTGCGTCACTGCTGCTACBC 550
QY 609 AGCGTTACTCCGCAATCCGATTCGAGATCTTGGAAAGCCATACCAATTCCTTAACAACA 668
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 551 GCGGCCACTCGCGCGTCCGCTGCTGCTGCTGAGGCGCTCAACCAATTCCTTAACAACA 610
QY 669 ACATTCACCTCTTTTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 728
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 611 GCATCAGCCCAATCGTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 670
QY 729 TATCATCATTCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 788
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 671 TCTCTATCATTCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 730
QY 789 GACAAGTCTCTCATTCGCGAAAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 845
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 731 AGCGCAAGGAGAGATCATCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 790
QY 846 TCGAGTTACAGCGGAAAGAGGCTAGCACTTTTAAAGGCGCGCGCGCGCGCGCGCGG 905
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 791 TCGTCTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 850
QY 906 TCGCTTCGATGCTTCTATTTGATGCTAATGCTACTTCTCTCTCTCTCTCTCTCTCTCT 965
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 851 TGGTGAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 910
QY 966 CGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1025
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 911 CTCTBACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 970
QY 1026 TGAAG--CATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1082
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 971 TCAAGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1030
QY 1083 GAAAGCGA--TTACGTCGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1139
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1031 GCGAGCTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1090
QY 1140 CAAAGCAAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1199
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1091 TCCGCGCGAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1150
QY 1200 TAAATCGGATCATACCAAAATATGATGAGAGGAAATATCAATTCCTCTCTCTCTCT 1259
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1151 ACATGATTCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1210
QY 1260 TGAATCGGATTCACCAAAATATGATGAGAGGAAATATCAATTCCTCTCTCTCTCTCT 1319
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1211 TCAATCGGATTCACCAAAATATGATGAGAGGAAATATCAATTCCTCTCTCTCTCTCT 1270
QY 1320 GAGTTTCGATTCAGCAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1379
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1271 CTAACAGGATTCAGCAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1330
QY 1380 AATTTCTGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1439

```

```

Db 1331 AGCTACCGGAGATGCTCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1387
QY 1440 GAGGTAACTTACTTGGACTACGAGGTTCAAAGGTGAGAAATATCCGCTTCTTACT 1499
Db 1388 CCGAGGACCCBCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1447
QY 1500 GTTCTGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1559
Db 1448 CTTCGAGAGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1507
QY 1560 ACAATCAAGAGCTTAATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1619
Db 1508 GCAACAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1567
QY 1620 ACATCTTAAACATCATCTGCTGACATCTATGATGCTCTATGCGCATTCATCTTAC 1679
Db 1568 AGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1627
QY 1680 GCCATTGGAAGAGAAATGAAATGCAAGTGAAGAGAAACAGGTAAGAGTCCGAGAAA 1739
Db 1628 GCGCGAGAGAGTTCAGATTCAAGAAAGCATTCGACCGCGCTTCTCTCTCTCTCTCT 1687
QY 1740 AGGTTCTTACTATGAGCGCTTAAAGGAGAGTTCGACCGCTCTCTCTCTCTCTCT 1793
Db 1688 AGCAGCTTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1741

```

RESULT 8

US-09-939-408A-20

```

: Sequence 20, Application US/09939408A
: Patent No. US20020102712A1
: GENERAL INFORMATION:
: APPLICANT: Yoshida, Roberta
: TITLE OF INVENTION: Phenylalanine Ammonia Lyase Polypeptide and
: TITLE OF INVENTION: Polynucleotide Sequences and Methods of Obtaining and
: TITLE OF INVENTION: Using Same
: FILE REFERENCE: 29479/50NCSA
: CURRENT APPLICATION NUMBER: US/09/939, 408A
: CURRENT FILING DATE: 2001-08-24
: PRIOR APPLICATION NUMBER: US 09/624,693
: PRIOR FILING DATE: 2000-07-24
: PRIOR APPLICATION NUMBER: PCT/US01/23270
: PRIOR FILING DATE: 2001-07-24
: NUMBER OF SEQ ID NOS: 30
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 20
: LENGTH: 2475
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: NAME/KEY: misc_difference
: LOCATION: (13) .. (2475)
: OTHER INFORMATION: n = A or C or G or T; "n" indicates no consensus at that posit
: OTHER INFORMATION: Description of Artificial Sequence: Consensus
: OTHER INFORMATION: Sequence of SEQ ID NOS: 12, 16, and 18
US-09-939-408A-20

```

```

Query Match      10.5% Score 256; DB 10; Length 2475;
Best Local Similarity 51.2%; Pred. No. 2.1e-52;
Matches 645; Conservative 0; Mismatches 600; Indels 15; Gaps 5;

```

```

QY 549 ACACACTTCACATTCAGCCACAGAGCCGCCATTCCTGCAATACACACCTCCCTCC 608
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 497 ACTGCTTCCGCTCGAGATGTTCTGCGGATATGATATCTGCGTCACTGCTGCTACBC 556
QY 609 AGCGTTACTCCGCAATCCGATTCGAGATCTTGGAAAGCCATACCAATTCCTTAACAACA 668
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 557 GCGGCCACTCGCGCGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 616
QY 669 ACATTCACCTCTTTTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 728
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


Db	713	AGGGCAAGGAGCAAGATCCTCTACGCGCGCGGAGCGAGCGATGGCGCTCTCTAAACGTCGAGCGCG	772
Gy	846	TGGAGTTATATCGTAAAGAAAGAGAGTAAATATTTATTAACAGATATAGAGATGAGAGATGGGGA	905
Db	773	TGCTGCTTCGCGCCCGCAAGAGAGGCTCTGCTGCTCTCAACGCGACCGCCGCTCTCAAGATTCGA	832
Gy	906	TGGCGTCGATCGTTCTATATTTGATGATGCTAAATCTACTTCGCTTGTGTGGCGAGTGGTATTCG	965
Db	834	TGGAAATATCTGCTCTGATAGAGAGATATATGATCTCTGGCTCTCTGGAGAGTGGCTCAAGG	892
Gy	966	CGATCTTCTGCTGAGGTATGCGAAGGCAAGCGCGGAGCTTACCGATACACTTTCGACAGCAAT	1035
Db	893	CGATACAGCGCTCGAAGCGCATATCTCTGCGCGACGCGGTCTGTTCGACCGCTTCCTTACGCG	952
Gy	1026	TGAATCTATATCTCTATTAATATATAGAGAGAGAGATATAGAGATATATTTGATG	1082
Db	953	TGACGCTGCGCGACCGCTAGATAG	1012
Gy	1083	GAA----CGGATTAAGTTCAGCG	1139
Db	1013	GAAATCTCTTTGCTGTATATATATAGAGAGAGAGATATAGAGATATAGAGAGAGATATG	1072
Gy	1140	CAAAACAAGATCTGTATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1196
Db	1073	TGCGCGAGAGAGCGTACCGCTCTCGCGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1132
Gy	1197	AGTATATCTGATATATCAAGTAAATATATATAGAGAGAGATATATTTCTATATATCAAGC	1256
Db	1133	ACGTCATTCG	1192
Gy	1257	CATTGATCGAGCTTTCGACAAACAAGCTTTACACGCGTGGTACCTTTCGACAGCAACCGCA	1316
Db	1193	CTCTCATCTGAGCTTATAGATATATATTTCTATATATAGAGAGATATATTTCTCTCTCTCTCT	1252
Gy	1317	TGCGAGTTTCGATGACAAACCGCGCTCTGCGCATCTGCTGCAATTCGAAATCTCATGTTTCG	1376
Db	1253	TGCGCAACACCTTCGACAAACAGCTGCGCTCTGCGCTCTGCGCGATGCGGCAAGCTCAACTTCA	1312
Gy	1377	CTCAATTTCTGAGTGTATTATATTTATATATATATATATATATATATATATATATATATATAT	1436
Db	1313	CGCAGCTCAGCGAGATATCTTAATCGCGCGGATGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCG	1372
Gy	1437	GTGAGCGTAACCCACTGTTTGGACTACCGCGTTCAGAGTGGAGAAATTCGCGATCGCTTCTT	1496
Db	1373	CGGAAAT	1429
Gy	1497	ACTGTTTGCAGCTTCAGTTTCTGCAAAATTCGATCGCAACCATGTTTCAAAAGCGCGGAC	1556
Db	1430	ACACCTGCGATTTGGCGACCTCTGCGCAACCGCTCTGAGAGCGATTCGACCGCGGCTGGA	1489
Gy	1557	AACAT	1616
Db	1490	TGGGGAACCTAGGCGCTTAACTGCTTGGCGCTATCTGCGCGCTCTCTGCGAGCGCGAGTCCA	1549
Gy	1617	TGCGAGATCTTAAAGATCATGCTGCGACATCACTATGCTGCTTATGCGCAATTCAGAT	1676
Db	1550	ACGAGCTCTTTCT	1609
Gy	1677	TACCG 1681	
Db	1610	TGCGC 1614	

RESULT 14
 US-09-765-873A-9
 : Sequence 9, Application US/09765873A
 : Patent No. US20010053847A1
 : GENERAL INFORMATION
 : APPLICANT: Tang, Xiao-Song
 : TITLE OF INVENTION: KETOPROTEIN OF FARA-HYDROXYCYNNAMIC ACID
 : FILE REFERENCE: BC1009 US CIP
 : CURRENT APPLICATION NUMBER: US/09/765,873A

[illegible]

Db 1253 TGGCCAAACACCATGAGAAAGACAGCGCTGCGCCGACAGATCCGCAAGCTCAACCTTCA 1312
QY 1377 CTCAATTTTGTAGTGTGGTTAAGATTTCTAACAACATGGATTTACCATGAACTCTCCG 1436
Db 1313 CGGAGCTCAGCGGATGCTCAAGCCGCGCATGACCCGCGCTCCGCTCCGCGCG 1372
QY 1437 GTGGAGTAAACCTTACTTGTGACTACGGGTGAAAGCTGAGAAATCCGCAATGCTTCTT 1496
Db 1373 CGGAA--GACCCCTGCTCTCTTACCATTTGAAGGCGCTGACATGCGCGCTGGGCGT 1429
QY 1497 ACTGTGTGAGCTTTCAGTTTCGCAATTCAGTCAACCAACATCTTCAAAAGCCGCAAC 1556
Db 1430 ACACCTGCGAGCTTGGACACCTCCGCAACCCCTGTGACGACGATCTCCAGCGCGCTGAGA 1489
QY 1557 AACCAATCAAGAGTTAATCTCTGCGATTAATTTCAAGCGGAAACCGGAGAGGAG 1616
Db 1490 TGGCGAAACGCGCGCTCAACCTCGCTTGGCTCATCTCGCTCTGTCGACAGACCGAGTCCA 1549
QY 1617 TCGACATCTTAAACATCATGCTGCGACATCACTTACGCTCTATGCGCAATTCATGAT 1676
Db 1550 ACGAGCTCTTCTCTCTCTCTGCGACACCACTTATTTGCTTTCGAAATCATGAGCT 1609
QY 1677 TACGC 1681
Db 1610 TCGCG 1614

RESULT 15
US-09-878-574-3252
; Sequence 3252. Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 3252
; LENGTH: 363
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: L1B3028-013-Q1-H1-C3
US-09-878-574-3252

Query Match 8.1%; Score 198.8; DB 10; Length 363;
Best Local Similarity 71.8%; Pred. No. 7.3e-39;
Matches 260; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 1418 TTACCAATTCATCTCCGCTGAGCTAACCTTGTGACTACGGGTCAAGGTGCA 1477
Db 2 TTGCTTTCAATCTCACTGCGACAGAAACCCAGCTTGATTTATGATTCAGAGAGCT 61
QY 1478 GAAATGCCCATGCTTCTTACTGCTTGTGAGCTTCACTTTCGCAAAATCCAGTCACCAAC 1537
Db 62 GAAATTTGCGCATGCGATCTTATTTGTGAACTTCAATATTTGGCGAATCCGCTGACAAGC 121
QY 1538 CATGTTCAAGGCGCGACACATCAAGAGCTTAATTTCTCTGGATTAAATTCAGCG 1597
Db 122 CAGGTGCAAAAGCGCGACACCAACCAAGATGTAACCTCTGCGCTGATTTCAATCA 181
QY 1598 ACGAAACCGGACGATCGACATCTTAAATCAATCAATGCTGAGATACCTTACCTGCT 1657
Db 182 AGAAGAAATTTATTAATTTATTTAAATTTATTAATTTATTTATTTATTTATTTATTT 241
QY 1658 CTATGCCAATTCATGATTTACCGATTTGAGAGAAACATGAATTCAGAGTGAAGAAC 1717
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 242 CTTTGGCAAGCCATTGACTTTGAGGCAATTTGAGAGAGAAATTTGAAGAACAGCGTCAAGAAC 301
QY 1718 ACGGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1777
Db 302 GTTGTAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 361
QY 1778 TC 1779
Db 362 TC 363
Search completed: March 29, 2003, 02:08:56
Job time : 213 secs

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 206-269-0565
 TELEFAX: 206-269-0563
 TELEX:
 INFORMATION FOR SEQ ID NO: 45:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 684 base pairs
 TYPE: nucleic acid
 STRANDNESS: single
 TOPOLOGY: linear
 US-08-975-316-45

Query Match 11.2%, Score 274.4, DB 2, Length 684.
 Best Local Similarity 70.7%; Pred. No. 4.7e-60;
 Matches 365, Conservative 0, Mismatches 151, Indels 0, Gaps 0;

```

OY 845 TTTCAGCTTACAGCCGAAGAGGCTAGCACTTTGTTAAGCGGACCGCGCTGGGCTCCGGG 904
    || || || || || || || || || || || || || || || || || || || || || ||
Db 169 TTTAAATTGAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 228
    || || || || || || || || || || || || || || || || || || || || || ||
OY 905 ATGGCTTCATGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 944
    || || || || || || || || || || || || || || || || || || || || || ||
Db 229 GTGGCCCGCTCCGCTCTGCTTTGACGCCACGCTGCTGCGCTGCTGCTGCTGCTGCTGCT 288
    || || || || || || || || || || || || || || || || || || || || || ||
OY 965 GATATTTTGTGAGTTATATAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1024
    || || || || || || || || || || || || || || || || || || || || || ||
Db 289 GCGCTCTCTCCGACGCTGATCCAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 348
    || || || || || || || || || || || || || || || || || || || || || ||
OY 1025 TTGAAGCATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1084
    || || || || || || || || || || || || || || || || || || || || || ||
Db 349 TTGAAGCATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 408
    || || || || || || || || || || || || || || || || || || || || || ||
OY 1085 AGCGATTACGTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1144
    || || || || || || || || || || || || || || || || || || || || || ||
Db 409 AGCGATTACGTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 468
    || || || || || || || || || || || || || || || || || || || || || ||
OY 1145 GATATTTTGTGAGTTATATAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1204
    || || || || || || || || || || || || || || || || || || || || || ||
Db 469 CAAGACCGCTTACGCTCTGCGCAACATCGCGCAACATCGCTGCGCGCGCGCGCGCGCG 528
    || || || || || || || || || || || || || || || || || || || || || ||
OY 1205 CGATCATCAGCAAGCAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1264
    || || || || || || || || || || || || || || || || || || || || || ||
Db 529 GCGCTCTCTCCGACGCTGATCCAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 588
    || || || || || || || || || || || || || || || || || || || || || ||
OY 1265 GACCTTTCAGCAAGCAAGCTTTAGACGCTGTAACCTTCAGCAAGCAAGCAAGCTT 1324
    || || || || || || || || || || || || || || || || || || || || || ||
Db 589 GATGCTCCAGGCAATGCGCTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 648
    || || || || || || || || || || || || || || || || || || || || || ||
OY 1325 TCCATGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
    || || || || || || || || || || || || || || || || || || || || || ||
Db 649 TCCATGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 684
    || || || || || || || || || || || || || || || || || || || || || ||

```

RESULT 7

US-09-615-192A-45

Sequence 45, Application US/09615192A

Patent No. 6410718

GENERAL INFORMATION:

APPLICANT: Hloksberg, Leonard N.

APPLICANT: Hloksberg, Leonard N.

TITLE OF INVENTION: Materials and Methods for the

TITLE OF INVENTION: Modification of Plant Lipin Content

FILE REFERENCE: 11000.1003040

CURRENT APPLICATION NUMBER: US/09/615, 192A

CURRENT FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 08/975,316

PRIOR FILING DATE: 1997-11-21

PRIOR APPLICATION NUMBER: US 08/713,000

PRIOR FILING DATE: 1996-09-11

PRIOR APPLICATION NUMBER: US 09/169,789

PRIOR FILING DATE: 1998-10-09

NUMBER OF SEQ ID NOS: 405

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 45
 LENGTH: 684
 TYPE: DNA
 ORGANISM: Pinus radiata
 US-09-615-192A-45

Query Match 11.2%, Score 274.4, DB 4, Length 684.
 Best Local Similarity 70.7%; Pred. No. 4.7e-60;
 Matches 365, Conservative 0, Mismatches 151, Indels 0, Gaps 0;

```

OY 845 TTTCAGCTTACAGCCGAAGAGGCTAGCACTTTGTTAAGCGGACCGCGCTGGGCTCCGGG 904
    || || || || || || || || || || || || || || || || || || || || || ||
Db 169 TTTAAATTGAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 228
    || || || || || || || || || || || || || || || || || || || || || ||
OY 905 ATGGCTTCATGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 944
    || || || || || || || || || || || || || || || || || || || || || ||
Db 229 GTGGCCCGCTCCGCTCTGCTTTGACGCCACGCTGCTGCGCTGCTGCTGCTGCTGCTGCT 288
    || || || || || || || || || || || || || || || || || || || || || ||
OY 965 GATATTTTGTGAGTTATATAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1024
    || || || || || || || || || || || || || || || || || || || || || ||
Db 289 GCGCTCTCTCCGACGCTGATCCAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 348
    || || || || || || || || || || || || || || || || || || || || || ||
OY 1025 TTGAAGCATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1084
    || || || || || || || || || || || || || || || || || || || || || ||
Db 349 TTGAAGCATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 408
    || || || || || || || || || || || || || || || || || || || || || ||
OY 1085 AGCGATTACGTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1144
    || || || || || || || || || || || || || || || || || || || || || ||
Db 409 AGCGATTACGTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 468
    || || || || || || || || || || || || || || || || || || || || || ||
OY 1145 GATATTTTGTGAGTTATATAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1204
    || || || || || || || || || || || || || || || || || || || || || ||
Db 469 CAAGACCGCTTACGCTCTGCGCAACATCGCGCAACATCGCTGCGCGCGCGCGCGCGCG 528
    || || || || || || || || || || || || || || || || || || || || || ||
OY 1205 CGATCATCAGCAAGCAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1264
    || || || || || || || || || || || || || || || || || || || || || ||
Db 529 GCGCTCTCTCCGACGCTGATCCAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 588
    || || || || || || || || || || || || || || || || || || || || || ||
OY 1265 GACCTTTCAGCAAGCAAGCTTTAGACGCTGTAACCTTCAGCAAGCAAGCAAGCTT 1324
    || || || || || || || || || || || || || || || || || || || || || ||
Db 589 GATGCTCCAGGCAATGCGCTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 648
    || || || || || || || || || || || || || || || || || || || || || ||
OY 1325 TCCATGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
    || || || || || || || || || || || || || || || || || || || || || ||
Db 649 TCCATGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 684
    || || || || || || || || || || || || || || || || || || || || || ||

```

RESULT 8

US-09-624-693A-20

Sequence 20, Application US/09624693A

Patent No. 6355468

GENERAL INFORMATION:

APPLICANT: Yoshida, Roberta

APPLICANT: Yoshida, Roberta

TITLE OF INVENTION: Phenylalanine Ammonia Lyase Polypeptide and

TITLE OF INVENTION: Polyaminoacid Sequences and Methods of Obtaining and

FILE REFERENCE: 29479/500NSC

CURRENT APPLICATION NUMBER: US/09/624,693A

CURRENT FILING DATE: 2000-07-24

NUMBER OF SEQ ID NOS: 25

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 20

LENGTH: 2475

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

NAME/KEY: misc difference

LOCATION: (13, 34, 46, 49, 51, 57, 59, 68, 69, 73, 75 - 77, 79, 82, 84,

OTHER INFORMATION: Description of Artificial Sequence: Consensus

OTHER INFORMATION: Sequence of SEQ ID NOS: 12, 16, and 18

Patent No. 6355468

NORTH OF SEABEARD, CO
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA

XX
DR M11; 2002-435152/46.
DR P-PSDB; AA097130.
PT Novel polynucleotide useful in preparing plant expression constructs to
PT modify the response of a plant cell to wounding, damage or other injury
PT from disease causing organisms or from plant pests -
XX
XX
PS Claim 1; Fig 5; 46pp. English.
XX
XX The present invention relates to the isolation of novel lettuce
CC (lactuca sativa) phenylalanine ammonia-lyase (LsPAL) enzymes, and the
CC polynucleotide sequences encoding them. An antibody that binds to
CC a PAL enzyme is useful for measuring the relative amount of PAL levels
CC in a tissue. The polynucleotide sequences encoding PAL are useful in
CC the preparation of plant expression constructs for modifying features
CC of the response of a plant cell to wounding, damage or other
CC injury from disease. Disruption or down-regulation of PAL activity
CC can be used to reduce the browning response to wounding in vegetable
CC crops. The present sequence encodes lettuce LsPAL enzyme.
XX
50 Sequence 2442 BP; 693 A; 553 C; 592 G; 604 T; 0 other;

Query Match	100.0%;	Score 2442;	DB 24;	Length 2442;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2442;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

[illegible]

Dd	661	TACACACATCACCCCTGTTTACCCTTCGGTGGACCACTACCGGCTCCGGTGA	720
Qy	721	TGTCCATTATCATATGAGCGAGCCTTTAACCGAGCGCCGCAATCCCAAGCGGTGG	780
Dd	721	TGTCCATTATCATATGAGCGAGCCTTTAACCGAGCGCCGCACTCCAAAGCGGTGG	780
Qy	781	CCCCACCGGAAATGCTTCATATGCGCAAAAAGGCTTGGTGGATGCGAGTGAAGGTGG	840
Dd	781	CCCCACCGGAAATGCTTCATATGCGCAAAAAGGCTTGGTGGATGCGAGTGAAGGTGG	840
Qy	841	GTTCTTCAGATTATATCGGAAAGAGAGATATGATTTTAAACGAGACGAGCGTGGGCTC	900
Dd	841	GTTCTTCAGATTATATCGGAAAGAGAGATATGATTTTAAACGAGACGAGCGTGGGCTC	900
Qy	901	CGGATGCGCTCCATGCTCTATTATGATATGATGACTGGCTGTGTGGAAAGTGT	960
Dd	901	CGGATGCGCTCCATGCTCTATTATGATATGATGACTGGCTGTGTGGAAAGTGT	960
Qy	961	ATCGGCATCTCCCTAGGTTATGCAAGGGAAGCGGAGTTTACCGATCATCTGACACA	1020
Dd	961	ATCGGCATCTCCCTAGGTTATGCAAGGGAAGCGGAGTTTACCGATCATCTGACACA	1020
Qy	1021	CAATTTGAGCATACCTTGGTAAATCGAGCGGTGATCATGAGTATATTTTGA	1080
Dd	1021	CAATTTGAGCATACCTTGGTAAATCGAGCGGTGATCATGAGTATATTTTGA	1080
Qy	1081	CGGAATATATTACTTAAAGTATGAGAAAGATTTATATGATGATGCGGTATACGAAAC	1140
Dd	1081	CGGAATGATTTACGTCAAAGCGCGCGCAAAAGTCCACGAATATGACCGCTTACAGAAAC	1140
Qy	1141	AAACCAAGATGTTATGCTCTCCGTACATCCCAATGGCTCGGACCTCAATGCAAT	1200
Dd	1141	AAACCAAGATGTTATGCTCTCCGTACATCTCCCAATGGCTCGGACCTCAATGCAAT	1200
Qy	1201	AATCGATCATCAACCAAAATGATCGAGAGGGAATCATCTCGTCACAGCAACCAT	1260
Dd	1201	AATCGATCATCAACCAAAATGATCGAGAGGGAATCATCTCGTCACAGCAACCAT	1260
Qy	1261	GATGAGCTTCCGAAACCAAGCTTTACAGGTGTAACTTCCCAAGAACCCCAATGG	1320
Dd	1261	GATGAGCTTCCGAAACCAAGCTTTACAGGTGTAACTTCCCAAGAACCCCAATGG	1320
Qy	1321	AGTTTCCATGCAACAACCGCTCCGCAATGCTGATATCGGAATATCATCTTCGCTCA	1380
Dd	1321	AGTTTCCATGCAACAACCGCTCCGCAATGCTGATATCGGAATATCATCTTCGCTCA	1380
Qy	1381	ATTTCCTAGCTGTTTAAGATTTTCTACAAATGATTTTCAATGATTTCTTCGGTGG	1440
Dd	1381	ATTTCCTAGCTGTTTAAGATTTTCTACAAATGATTTTCAATGATTTCTTCGGTGG	1440
Qy	1441	ACGTAACTTTAGTTTGAATTCAGAGGTTCAAAAGTGGAGAAATCGGATTTTACTG	1500
Dd	1441	ACGTAACTTTAGTTTGAATTCAGAGGTTCAAAAGTGGAGAAATCGGATTTTACTG	1500
Qy	1501	TTTCAGACTTCAGTTTTCGCAAAATCGAGTACCAACATGTTTCAAGCGCGCAACACA	1560
Dd	1501	TTTCAGACTTCAGTTTTCGCAAAATCGAGTACCAACATGTTTCAAGCGCGCAACACA	1560
Qy	1561	CAATCAAGAGTTAATTTCTTGATTTATTTGAGGAGGAAACCTGAGAGAGTGTGA	1620
Dd	1561	CAATCAAGAGTTAATTTCTTGATTTATTTTCAAGGAGGAAACCTGAGAGTGTGA	1620
Qy	1621	CATTTTAAATATATGTCGTGGATATTTAGTGGATTTATATATATCATCGATTACG	1680
Dd	1621	CATTTTAAATATATGTCGTGGATATTTAGTGGATTTATATATATCATCGATTACG	1680
Qy	1681	CCATTGGAAGACAGCAATGAAATCGACAGTGAAGAACCGGTATACCGCAAGTCCGCAAAA	1740
Dd	1681	CCATTGGAAGACAGCAATGAAATCGACAGTGAAGAACCGGTATACCGCAAGTCCGCAAAA	1740
Qy	1741	GATCTTAACTATGAGATGAGATATGAGATTTGAGATTTGAGATGATGT	1800
Dd	1741	GATCTTAACTATGAGATGAGATATGAGATTTGAGATTTGAGATGATGT	1800

[illegible]

QY 1801 CTCTGCTGTTGATGCTTAATGCTTTTCTTATATGAGAGGTTTGAGGAGC 1860
 DB 1801 CTTCTGCTGTTGATGCTTAATGCTTTTCTTATATGAGAGGTTTGAGGAGC 1860
 QY 1861 ATATGATTAATGCTTAATGCTTTTCTTATATGAGAGGTTTGAGGAGC 1920
 DB 1861 ATATGATTAATGCTTAATGCTTTTCTTATATGAGAGGTTTGAGGAGC 1920
 QY 1921 AACGAGACAGACGTAACGCTTTTCTTATATGAGAGGTTTGAGGAGC 1980
 DB 1921 AACGAGACAGACGTAACGCTTTTCTTATATGAGAGGTTTGAGGAGC 1980
 QY 1981 GAAAGTCTTTTCTTATATGAGAGGTTTGAGGAGGTTTGAGGAGC 2040
 DB 1981 GAAAGTCTTTTCTTATATGAGAGGTTTGAGGAGGTTTGAGGAGC 2040
 QY 2041 GTTATTTTAAATGATTAATGCTTTTCTTATATGAGAGGTTTGAGGAGC 2100
 DB 2041 GTTATTTTAAATGATTAATGCTTTTCTTATATGAGAGGTTTGAGGAGC 2100
 QY 2101 GCACTGCTGAGACGCTTTTCTTATATGAGAGGTTTGAGGAGC 2160
 DB 2101 GCACTGCTGAGACGCTTTTCTTATATGAGAGGTTTGAGGAGC 2160
 QY 2161 GAGAGTCTTTTCTTATATGAGAGGTTTGAGGAGGTTTGAGGAGC 2220
 DB 2161 GAGAGTCTTTTCTTATATGAGAGGTTTGAGGAGGTTTGAGGAGC 2220
 QY 2221 AGGATGATTAATGCTTTTCTTATATGAGAGGTTTGAGGAGC 2280
 DB 2221 AGGATGATTAATGCTTTTCTTATATGAGAGGTTTGAGGAGC 2280
 QY 2281 TGTATTTTAAATGATTAATGCTTTTCTTATATGAGAGGTTTGAGGAGC 2340
 DB 2281 TGTATTTTAAATGATTAATGCTTTTCTTATATGAGAGGTTTGAGGAGC 2340
 QY 2341 CATGATTAATGCTTTTCTTATATGAGAGGTTTGAGGAGGTTTGAGGAGC 2400
 DB 2341 CATGATTAATGCTTTTCTTATATGAGAGGTTTGAGGAGGTTTGAGGAGC 2400
 QY 2401 ATATTTTCTTATATGAGAGGTTTGAGGAGGTTTGAGGAGC 2442
 DB 2401 ATATTTTCTTATATGAGAGGTTTGAGGAGGTTTGAGGAGC 2442

RESULT 2
 ABR50680
 ID ABR50680 standard; cDNA: 2380 BP.
 XX
 AC ABR50680:
 DT 13-AUG-2002 (first entry)
 XX
 DE cDNA encoding lettuce LsPAL2 enzyme.
 XX
 KW lettuce; phenylalanine ammonia-lyase, plant wounding; plant damage;
 KW plant disease; plant browning; vegetable crop; plant; LsPAL2;
 KW gene; ss.
 XX
 OS Lactuca salvia.
 XX
 FH Key location/Organisms
 FH CDS 77..2218
 FT /tag- a
 FT /product- "LsPAL2"
 FT /transl_except- (pos:135..145, aa:Val-Gly)
 XX
 FN W0200226028.A2.
 PD 04-APR-2002.
 XX
 PF 26-SEP-2001; 2001W0-NS30192.

XX 26-SEP-2000; 2000NS-235956P
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Salveit M, Campos R, Nonogaki H, Sussow T;
 DB WPI: 2002-435152/46.
 DB P-PSDB; AA097131.
 PT Novel polynucleotide useful in preparing plant expression constructs to
 PT modify the response of a plant cell to wounding, damage or other injury
 PT from disease causing organisms or from plant pests
 PS Claim 2; Fig 7; 46pp; English.
 CC The present invention relates to the isolation of novel lettuce
 CC (Lactuca salvia) phenylalanine ammonia lyase (LsPAL) enzymes, and the
 CC polynucleotide sequences encoding them. An antibody that binds to
 CC a PAL enzyme is useful for measuring the relative amount of PAL levels
 CC in a tissue. The polynucleotide sequences encoding PAL are useful in
 CC the preparation of plant expression constructs for modifying features
 CC of the response of a plant cell to wounding, damage or other
 CC injury from disease. Disruption or down-regulation of PAL activity
 CC can be used to reduce the browning response to wounding in vegetable
 CC crops. The present sequence encodes lettuce LsPAL2 enzyme.
 CC
 SQ Sequence 2380 BP; 663 A, 548 C, 603 G, 566 T; 0 other.
 Query Match 65.7%, Score 2603.8, Lb 24, Length 2380;
 Best Local Similarity 85.5%; Pred. No. 4e-278;
 Matches 1812; Conservative 0; Mismatches 296; Indels 12; Gaps 2;
 QY 169 GATTTTAAATGATTAATGCTTTTCTTATATGAGAGGTTTGAGGAGC 228
 DB 145 GATTTTAAATGATTAATGCTTTTCTTATATGAGAGGTTTGAGGAGC 204
 QY 229 GAAATGATTAATGCTTTTCTTATATGAGAGGTTTGAGGAGGTTTGAGGAGC 2340
 DB 205 GAAATGATTAATGCTTTTCTTATATGAGAGGTTTGAGGAGGTTTGAGGAGC 2340
 QY 289 AGTTTCTTATATGAGAGGTTTGAGGAGGTTTGAGGAGGTTTGAGGAGC 348
 DB 265 GTTATTTTAAATGATTAATGCTTTTCTTATATGAGAGGTTTGAGGAGC 321
 QY 349 GTTATTTTAAATGATTAATGCTTTTCTTATATGAGAGGTTTGAGGAGC 408
 DB 322 GTTATTTTAAATGATTAATGCTTTTCTTATATGAGAGGTTTGAGGAGC 381
 QY 409 TAAATGATTAATGCTTTTCTTATATGAGAGGTTTGAGGAGGTTTGAGGAGC 468
 DB 382 TAAATGATTAATGCTTTTCTTATATGAGAGGTTTGAGGAGGTTTGAGGAGC 441
 QY 469 TAAATGATTAATGCTTTTCTTATATGAGAGGTTTGAGGAGGTTTGAGGAGC 528
 DB 442 GAAATGATTAATGCTTTTCTTATATGAGAGGTTTGAGGAGGTTTGAGGAGC 501
 QY 529 GAAATGATTAATGCTTTTCTTATATGAGAGGTTTGAGGAGGTTTGAGGAGC 588
 DB 502 TAAATGATTAATGCTTTTCTTATATGAGAGGTTTGAGGAGGTTTGAGGAGC 561
 QY 589 GAAATGATTAATGCTTTTCTTATATGAGAGGTTTGAGGAGGTTTGAGGAGC 648
 DB 562 GAAATGATTAATGCTTTTCTTATATGAGAGGTTTGAGGAGGTTTGAGGAGC 621
 QY 649 GAAATGATTAATGCTTTTCTTATATGAGAGGTTTGAGGAGGTTTGAGGAGC 708
 DB 622 GAAATGATTAATGCTTTTCTTATATGAGAGGTTTGAGGAGGTTTGAGGAGC 681
 QY 709 GAAATGATTAATGCTTTTCTTATATGAGAGGTTTGAGGAGGTTTGAGGAGC 768
 DB 682 GAAATGATTAATGCTTTTCTTATATGAGAGGTTTGAGGAGGTTTGAGGAGC 741

QY 769 CAAAGCCGTCGCCCCACCGAGAGTCCTCATATGCCGAAAGAGCCCTTCGTCGACCGG 828
 Db 742 CAAAGCCGTCGCCCCACCGAGAGTCCTCATATGCCGAAAGAGCCCTTCGTCGACCGG 801
 QY 829 AGTGTAGGCTGGGTTCTTGAGTTACAGCCGAAAGAGAGAGTAAATTTGTAATTTGAC 888
 Db 802 AGTGTAGGCTGGGTTCTTGAGTTACAGCCGAAAGAGAGAGAGTAAATTTGTAATTTGAC 861
 QY 889 CAGCGTGGGCTGGGATGGCTTCATGTTCTTATTTGATCTGATGTTGTTGTTGTTGTT 948
 Db 862 CAGCGTGGGATGGGATGGCTTCATGTTCTTATTTGATCTGATGTTGTTGTTGTTGTTGTT 921
 QY 949 GTGCGAAGTGTATCGCGCATCTTCGTGAGGTTATGCGAAGGAGCGCGGATTTACCGA 1008
 Db 922 GTGCGAAGTGTATCGCGCATCTTCGTGAGGTTATGCGAAGGAGCGCGGATTTACTGA 981
 QY 1009 TCAGCTTGAATCAATTAATTAAGGATCACTCGGTGTAATTCGAGCGCGCGGATCATGA 1068
 Db 982 TCAGCTTGAATCAATTAATTAAGGATCACTCGGTGTAATTCGAGCGCGCGGATCATGA 1041
 QY 1069 GTATATTTTGAGCGAGAGCATTAAGTCAGGCGGCGTAAAGGTTGTCAGMAATGAGACC 1128
 Db 1042 GTATATTTTGAGCGAGAGCATTAAGTCAGGCGGCGGCGTAAAGGTTGTCAGMAATGAGACC 1101
 QY 1129 GTTACGAGAAACCAACAAAGATCGTTATGCTCTCGCTACATCTCCCAATGCTCGGACC 1188
 Db 1102 GTTACGAGAAACCAACAAAGATCGTTATGCTCTCGCTACATCTCCCAATGCTCGGACC 1161
 QY 1189 TCAGATCGAATATGATGATATATCAACAAATGATGAGGAGGAAATTCATCTCGTCA 1248
 Db 1162 TCAGATCGAATATGATGATATATCAACAAATGATGAGGAGGAAATTCATCTCGTCA 1221
 QY 1249 CGACAAACCCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1308
 Db 1222 CGACAAACCCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1281
 QY 1309 AACCCCAATCGGAGTTTCATGAGCAACACCCGCTCTCCCATCTTCGCAATCGGAAACT 1368
 Db 1282 AACCCCAATCGGAGTTTCATGAGCAACACCCGCTCTCCCATCTTCGCAATCGGAAACT 1241
 QY 1369 CATGTGCGTCAATTTCTGAGCTGGTTAGCATTTCTACAAACATGGATTACATCGAA 1428
 Db 1342 CATGTGCGTCAATTTCTGAGCTGGTTAGCATTTCTACAAACATGGATTACATCGAA 1401
 QY 1429 TCTTTCGAGTGAACGTAAATAGTTTGAATAGAGGTTTAAAGTGAAGAAATGCGCAT 1488
 Db 1402 TCTTTCGAGTGAACGTAAATAGTTTGAATAGAGGTTTAAAGTGAAGAAATGCGCAT 1461
 QY 1489 GCGTCTTACTGTTCTGAGCTTCATCTTCGCAAAATGCAATCGCAACCATCTTCAAG 1548
 Db 1462 GCGTCTTACTGTTCTGAGCTTCATCTTCGCAAAATGCAATCGCAACCATCTTCAAG 1521
 QY 1549 CGCGCGAATATGCAATGATTAATTTCTCGGATTAAATTCAGCGAGAAACCGC 1608
 Db 1522 CGCGCGAATATGCAATGATTAATTTCTCGGATTAAATTCAGCGAGAAACCGC 1581
 QY 1609 AAGAGAGTGAATCTTAATAATCTATGTCGACATCTTACTGATCTTATGTCATG 1668
 Db 1582 AAGAGAGTGAATCTTAATAATCTATGTCGACATCTTACTGATCTTATGTCATG 1641
 QY 1669 CATGCAATTAACGCAATTTGAGAGCAACATGCAATGCAAGAGAGAGCAACATGTAAGCTA 1728
 Db 1642 CATGCAATTAACGCAATTTGAGAGCAACATGCAATGCAAGAGAGAGCAACATGTAAGCTA 1701
 QY 1729 AATGCGAAGAAAGGTTCTTATGAGGAGTAAAGTGAATGTAATGTAATGTAATGTAATG 1788
 Db 1702 AATGCGAAGAAAGGTTCTTATGAGGAGTAAAGTGAATGTAATGTAATGTAATGTAATG 1761
 QY 1789 CGAGAGAAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1848
 Db 1762 CGAGAGAAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1821
 QY 1849 TTGAGCGCGACATACCATTAATGCAAGAGCTCGGACAGGTTCTGCTGACACCGCTCT 1908

Db 1822 TTGAGCGCGACACCTTACCATTTGATGAGAAAGCTCCGACAGGTTATGCTGACACCGCAT 1881
 QY 1909 AAAAATAATGAG 1968
 Db 1882 AAACAAAC-----CAAAATGACCGCGAGAACTTCATCTTCCAAAAGATGATGAT 1932
 QY 1969 CGAAGAGAAATGAG 2028
 Db 1933 CGAAGAGAAATGAG 1992
 QY 2029 GATGATACATTTGCTGATTTGCAACAGAGATTAAGCTTGCAGATCGTACCGTTGATAG 2088
 Db 1993 GATGATACATTTGCTGATTTGCAACAGAGATTAAGCTTGCAGATCGTACCGTTGATAG 2052
 QY 2089 GTTGTAG 2148
 Db 2053 GTTGTAG 2112
 QY 2149 AGAGAGATTTGCAAG 2208
 Db 2113 AGAGAGATTTGCAAG 2172
 QY 2209 GAGAGTCTTTGAG 2268
 Db 2173 GAGAGTCTTTGAG 2232
 QY 2269 AACCGTTGAAATTTGATTTG 2288
 Db 2233 AATACTTTTATGAGCTTTTG 2252

RESULT 3
 AAC51556
 ID AAC51556 standard, DNA, 2400 BP.
 XX
 AC AAC51556;
 XX
 Dt 18-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 68947.
 XX
 KW Hybridisation assay; genetic mapping; gene expression control;
 KW protein identification; signal transduction pathway;
 KW metabolic pathway; promoter; termination sequence; ss.
 OS Arabidopsis thaliana.
 ON
 PN EP103405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127452.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486-
PR 06-MAY-1999; 99US-0132487
PR 07-MAY-1999; 99US-0132487
PR 11-MAY-1999; 99US-0132486
PR 14-MAY-1999; 99US-0134218
PR 14-MAY-1999; 99US-0134219
PR 14-MAY-1999; 99US-0134220
PR 14-MAY-1999; 99US-0134221
PR 14-MAY-1999; 99US-0134221
PR 18-MAY-1999; 99US-0134768
PR 19-MAY-1999; 99US-0134941
PR 20-MAY-1999; 99US-0135124
PR 21-MAY-1999; 99US-0135153
PR 24-MAY-1999; 99US-0135629
PR 25-MAY-1999; 99US-0136021
PR 27-MAY-1999; 99US-0136392
PR 28-MAY-1999; 99US-0136782
PR 01-JUN-1999; 99US-0137222
PR 01-JUN-1999; 99US-0137528
PR 04-JUN-1999; 99US-0137502
PR 07-JUN-1999; 99US-0137724
PR 08-JUN-1999; 99US-0138094
PR 10-JUN-1999; 99US-0138540
PR 10-JUN-1999; 99US-0138847
PR 14-JUN-1999; 99US-0139119
PR 16-JUN-1999; 99US-0139452
PR 16-JUN-1999; 99US-0139453
PR 17-JUN-1999; 99US-0139492
PR 18-JUN-1999; 99US-0139454
PR 18-JUN-1999; 99US-0139455
PR 18-JUN-1999; 99US-0139456
PR 18-JUN-1999; 99US-0139457
PR 18-JUN-1999; 99US-0139458
PR 18-JUN-1999; 99US-0139459
PR 18-JUN-1999; 99US-0139460
PR 18-JUN-1999; 99US-0139461
PR 18-JUN-1999; 99US-0139462
PR 18-JUN-1999; 99US-0139463
PR 18-JUN-1999; 99US-0139760
PR 18-JUN-1999; 99US-0139763
PR 21-JUN-1999; 99US-0139817
PR 22-JUN-1999; 99US-0139899
PR 23-JUN-1999; 99US-0140353
PR 23-JUN-1999; 99US-0140354
PR 24-JUN-1999; 99US-0140645
PR 28-JUN-1999; 99US-0140823
PR 29-JUN-1999; 99US-0140991
PR 30-JUN-1999; 99US-0141287
PR 01-JUL-1999; 99US-0141842
PR 01-JUL-1999; 99US-0142154
PR 02-JUL-1999; 99US-0142055
PR 06-JUL-1999; 99US-0142190
PR 08-JUL-1999; 99US-0142803
PR 09-JUL-1999; 99US-0142920
PR 12-JUL-1999; 99US-0142977
PR 13-JUL-1999; 99US-0143542
PR 14-JUL-1999; 99US-0143624
PR 15-JUL-1999; 99US-0144085
PR 16-JUL-1999; 99US-0144085
PR 16-JUL-1999; 99US-0144086
PR 19-JUL-1999; 99US-0144125
PR 19-JUL-1999; 99US-0144131
PR 19-JUL-1999; 99US-0144132
PR 19-JUL-1999; 99US-0144133
PR 19-JUL-1999; 99US-0144134
PR 19-JUL-1999; 99US-0144134
PR 20-JUL-1999; 99US-0144335
PR 20-JUL-1999; 99US-0144352
PR 20-JUL-1999; 99US-0144352
PR 26-JUL-1999; 99US-0144632
PR 21-JUL-1999; 99US-0144684
PR 21-JUL-1999; 99US-0144684
PR 21-JUL-1999; 99US-0145086
PR 21-JUL-1999; 99US-0145088
PR 22-JUL-1999; 99US-0145085
PR 22-JUL-1999; 99US-0145087

PR 22-JUL-1999; 99US-0145089
PR 22-JUL-1999; 99US-0145192
PR 23-JUL-1999; 99US-0145145
PR 23-JUL-1999; 99US-0145218
PR 23-JUL-1999; 99US-0145224
PR 26-JUL-1999; 99US-0145476
PR 27-JUL-1999; 99US-0145913
PR 27-JUL-1999; 99US-0145918
PR 27-JUL-1999; 99US-0145919
PR 28-JUL-1999; 99US-0145951
PR 02-AUG-1999; 99US-0146386
PR 02-AUG-1999; 99US-0146388
PR 02-AUG-1999; 99US-0146389
PR 03-AUG-1999; 99US-0147038
PR 04-AUG-1999; 99US-0147204
PR 04-AUG-1999; 99US-0147302
PR 05-AUG-1999; 99US-0147192
PR 05-AUG-1999; 99US-0147260
PR 06-AUG-1999; 99US-0147303
PR 06-AUG-1999; 99US-0147416
PR 09-AUG-1999; 99US-0147493
PR 09-AUG-1999; 99US-0147935
PR 10-AUG-1999; 99US-0148171
PR 11-AUG-1999; 99US-0148319
PR 12-AUG-1999; 99US-0148341
PR 13-AUG-1999; 99US-0148565
PR 13-AUG-1999; 99US-0148684
PR 16-AUG-1999; 99US-0149368
PR 17-AUG-1999; 99US-0149175
PR 18-AUG-1999; 99US-0149426
PR 20-AUG-1999; 99US-0149722
PR 20-AUG-1999; 99US-0149723
PR 20-AUG-1999; 99US-0149929
PR 23-AUG-1999; 99US-0149902
PR 23-AUG-1999; 99US-0149930
PR 25-AUG-1999; 99US-0150566
PR 26-AUG-1999; 99US-0150884
PR 27-AUG-1999; 99US-0151065
PR 27-AUG-1999; 99US-0151066
PR 30-AUG-1999; 99US-0151080
PR 31-AUG-1999; 99US-0151303
PR 31-AUG-1999; 99US-0151338
PR 01-SEP-1999; 99US-0151430
PR 07-SEP-1999; 99US-0152363
PR 10-SEP-1999; 99US-0153070
PR 13-SEP-1999; 99US-0153758
PR 15-SEP-1999; 99US-0154039
PR 16-SEP-1999; 99US-0154039
PR 20-SEP-1999; 99US-0154479
PR 22-SEP-1999; 99US-0155139
PR 23-SEP-1999; 99US-0155139
PR 24-SEP-1999; 99US-0155486
PR 28-SEP-1999; 99US-0155659
PR 29-SEP-1999; 99US-0156596
PR 04-OCT-1999; 99US-0157117
PR 05-OCT-1999; 99US-0157753
PR 06-OCT-1999; 99US-0157865
PR 07-OCT-1999; 99US-0158029
PR 08-OCT-1999; 99US-0158332
PR 12-OCT-1999; 99US-0158369
PR 13-OCT-1999; 99US-0159293
PR 13-OCT-1999; 99US-0159294
PR 13-OCT-1999; 99US-0159295
PR 14-OCT-1999; 99US-0159329
PR 14-OCT-1999; 99US-0159330
PR 14-OCT-1999; 99US-0159331
PR 14-OCT-1999; 99US-0159637
PR 14-OCT-1999; 99US-0159638
PR 18-OCT-1999; 99US-0159584
PR 21-OCT-1999; 99US-0160741
PR 21-OCT-1999; 99US-0160767
PR 21-OCT-1999; 99US-0160768
PR 21-OCT-1999; 99US-0160770

```
PR 21-OCT-1999: 9905-0160814.
PR 21-OCT-1999: 9905-0160815.
PR 22-OCT-1999: 9905-0160980.
PR 22-OCT-1999: 9905-0160981.
PR 22-OCT-1999: 9905-0160989.
PR 25-OCT-1999: 9905-0161404.
PR 25-OCT-1999: 9905-0161405.
PR 25-OCT-1999: 9905-0161406.
PR 26-OCT-1999: 9905-0161359.
PR 26-OCT-1999: 9905-0161360.
PR 26-OCT-1999: 9905-0161361.
PR 28-OCT-1999: 9905-0161920.
PR 28-OCT-1999: 9905-0161922.
PR 28-OCT-1999: 9905-0161993.
PR 29-OCT-1999: 9905-0162142.

Query Match 49.8%; Score 1215.8, DB 21; Length 2400,
Best Local Similarity 72.8%, Prd. No. 124-208;
Matches 1584; Conservative 0; Mismatches 587; Indels 6; Gaps 1;

QY 140 AATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 199
DB 148 AATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 207
QY 200 GGTGTTAAAGAAATTAATGATAGATGAGAAATGATGATGAGAAATGAGAAATG 259
DB 208 CAATGCAAGAAAGTCAATTTAGATGAGAAATGAGAAATGAGAAATGAGAAATGAG 267
QY 260 GTGCTGAGCTGGAGAGAGAGAGAGCTTCTGAGTGGGAGGAGAGAGAGAGAGAG 319
DB 268 GTGCTGAGCTGGAGAGAGAGAGAGCTTCTGAGTGGGAGGAGAGAGAGAGAGAGAG 321
QY 320 AATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 379
DB 322 AATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 381
QY 380 AATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 439
DB 382 AATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 441
QY 440 GGTGTTAAAGAAATTAATGATAGATGAGAAATGATGATGAGAAATGAGAAATGAG 499
DB 442 GGTGTTAAAGAAATTAATGATAGATGAGAAATGATGATGAGAAATGAGAAATGAG 501
QY 500 AATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 559
DB 502 AATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 561
QY 560 GATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 619
DB 562 GATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 621
QY 620 GATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 679
DB 622 GATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 681
QY 680 TGTGTTAAAGAAATTAATGATAGATGAGAAATGATGATGAGAAATGAGAAATGAG 739
DB 682 TGTGTTAAAGAAATTAATGATAGATGAGAAATGATGATGAGAAATGAGAAATGAG 741
QY 740 GGTGTTAAAGAAATTAATGATAGATGAGAAATGATGATGAGAAATGAGAAATGAG 799
DB 742 GGTGTTAAAGAAATTAATGATAGATGAGAAATGATGATGAGAAATGAGAAATGAG 801
QY 800 AATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 859
DB 802 AATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 861
QY 860 AATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 919
DB 862 AATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 921
QY 920 CTATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 979

DB 922 CTATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 981
QY 980 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1039
DB 982 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1041
QY 1040 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1099
DB 1042 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1101
QY 1100 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1159
DB 1102 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1161
QY 1160 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1219
DB 1162 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1221
QY 1220 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1279
DB 1222 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1281
QY 1280 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1339
DB 1282 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1341
QY 1340 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1399
DB 1342 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1401
QY 1400 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1459
DB 1402 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1461
QY 1460 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1519
DB 1462 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1521
QY 1520 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1579
DB 1522 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1581
QY 1580 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1639
DB 1582 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1641
QY 1640 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1699
DB 1642 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1701
QY 1700 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1759
DB 1702 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1761
QY 1760 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1819
DB 1762 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1821
QY 1820 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1879
DB 1822 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1881
QY 1880 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1939
DB 1882 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1941
QY 1940 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 1999
DB 1942 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 2001
QY 2000 GTTATGAGTCTTAAATGAGTTGTGATCAAGATCCATTTGAGTGGGAGCTTGACGGAG 2059
```


[illegible]

PR	13-AUG-1999;	9905-0148684
PR	16-AUG-1999;	9905-0149175
PR	17-AUG-1999;	9905-0149186
PR	18-AUG-1999;	9905-0149426
PR	20-AUG-1999;	9905-0149722
PR	20-AUG-1999;	9905-0149723
PR	20-AUG-1999;	9905-0149929
PR	23-AUG-1999;	9905-0149902
PR	23-AUG-1999;	9905-0149930
PR	25-AUG-1999;	9905-0105066
PR	26-AUG-1999;	9905-0105084
PR	27-AUG-1999;	9905-0151065
PR	27-AUG-1999;	9905-0151066
PR	27-AUG-1999;	9905-0151080
PR	30-AUG-1999;	9905-0151303
PR	31-AUG-1999;	9905-0151338
PR	01-SEP-1999;	9905-0151340
PR	07-SEP-1999;	9905-0152363
PR	10-SEP-1999;	9905-0153557
PR	13-SEP-1999;	9905-0153758
PR	15-SEP-1999;	9905-0154019
PR	16-SEP-1999;	9905-0154038
PR	20-SEP-1999;	9905-0154779
PR	22-SEP-1999;	9905-0155139
PR	23-SEP-1999;	9905-0155585
PR	24-SEP-1999;	9905-0155659
PR	28-SEP-1999;	9905-0156458
PR	29-SEP-1999;	9905-0156596
PR	04-OCT-1999;	9905-0157117
PR	05-OCT-1999;	9905-0157753
PR	06-OCT-1999;	9905-0157865
PR	07-OCT-1999;	9905-0158029
PR	08-OCT-1999;	9905-0158232
PR	12-OCT-1999;	9905-0158369
PR	13-OCT-1999;	9905-0159294
PR	13-OCT-1999;	9905-0159294
PR	13-OCT-1999;	9905-0159295
PR	14-OCT-1999;	9905-0159329
PR	14-OCT-1999;	9905-0159330
PR	14-OCT-1999;	9905-0159331
PR	14-OCT-1999;	9905-0159637
PR	14-OCT-1999;	9905-0159684
PR	18-OCT-1999;	9905-0159684
PR	21-OCT-1999;	9905-0160741
PR	21-OCT-1999;	9905-0160767
PR	21-OCT-1999;	9905-0160768
PR	21-OCT-1999;	9905-0160770
PR	21-OCT-1999;	9905-0160814
PR	21-OCT-1999;	9905-0160815
PR	22-OCT-1999;	9905-0160815
PR	22-OCT-1999;	9905-0160881
PR	22-OCT-1999;	9905-0160981
PR	23-OCT-1999;	9905-0161004
PR	23-OCT-1999;	9905-0161405
PR	25-OCT-1999;	9905-0161406
PR	26-OCT-1999;	9905-0161359
PR	26-OCT-1999;	9905-0161360
PR	26-OCT-1999;	9905-0161361
PR	26-OCT-1999;	9905-0161362
PR	28-OCT-1999;	9905-0161920
PR	28-OCT-1999;	9905-0161923
PR	28-OCT-1999;	9905-0162142

Query Match	43.3%	Score 1200.6;	DB 21;	Length 2402;
Best Local Similarity	72.6%;	Pred. No. 6.2e-206;		
Matches 1581; Conservative	0;	Mismatches 589;	Indels 7;	Gaps 2

[illegible]

[illegible]

XX	25 FEB 2000	4000PR-0401439
XX		
FR	25 FEB-1999	9905-01231825
FR	03-MAR-1999	9905-0123180
PR	09-MAR-1999	9905-01235348
PR	23-MAR-1999	9905-0125758
PR	25-MAR-1999	9905-0126564
PR	29-MAR-1999	9905-0126785
PR	01-APR-1999	9905-0127462
PR	06-APR-1999	9905-0128234
PR	08-APR-1999	9905-0128714
PR	16-APR-1999	9905-0129645
PR	19-APR-1999	9905-0130077
PR	21-APR-1999	9905-0130549
PR	24-APR-1999	9905-0130440
PR	24-APR-1999	9905-0130891
PR	28-APR-1999	9905-0131449
PR	30-APR-1999	9905-0132048
PR	30-APR-1999	9905-0132407
PR	04-MAY-1999	9905-0132484
PR	06-MAY-1999	9905-0132485
PR	06-MAY-1999	9905-0132485
PR	06-MAY-1999	9905-0132487
PR	07-MAY-1999	9905-0132863
PR	11-MAY-1999	9905-0134256
PR	14-MAY-1999	9905-0134218
PR	14-MAY-1999	9905-0134219
PR	14-MAY-1999	9905-0134221
PR	14-MAY-1999	9905-0134370
PR	18-MAY-1999	9905-0134768
PR	19-MAY-1999	9905-0134941
PR	20-MAY-1999	9905-0135124
PR	21-MAY-1999	9905-0135253
PR	24-MAY-1999	9905-0135459
PR	25-MAY-1999	9905-0136021
PR	27-MAY-1999	9905-0136392
PR	28-MAY-1999	9905-0136782
PR	01-JUN-1999	9905-0137222
PR	03-JUN-1999	9905-0137428
PR	04-JUN-1999	9905-0137602
PR	07-JUN-1999	9905-0137724
PR	08-JUN-1999	9905-0138094
PR	10-JUN-1999	9905-0138540
PR	14-JUN-1999	9905-0138647
PR	16-JUN-1999	9905-0139452
PR	16-JUN-1999	9905-0139493
PR	17-JUN-1999	9905-0139492
PR	18-JUN-1999	9905-0139454
PR	18-JUN-1999	9905-0139455
PR	18-JUN-1999	9905-0139456
PR	18-JUN-1999	9905-0139457
PR	18-JUN-1999	9905-0139458
PR	18-JUN-1999	9905-0139459
PR	18-JUN-1999	9905-0139460
PR	23-JUN-1999	9905-0139461
PR	24-JUN-1999	9905-0140354
PR	24-JUN-1999	9905-0140695
PR	29-JUN-1999	9905-0140992
PR	30-JUN-1999	9905-0141287
PR	01-JUL-1999	9905-0141464
PR	01-JUL-1999	9905-0142154
PR	02-JUL-1999	9905-0142390
PR	06-JUL-1999	9905-0142595

xx
Sequence 5291 BP; 1512 A; 1061 C; 971 G; 1747 T; 0 other;

Query Match 41.2%; Score 1005.2; DB 11, Length 5291;
Best Local Similarity 73.3%; Pred. No. 6.6e-171;
Matches 1286; Conservative 0; Mismatches 468; Indels 0; Gaps 0;

QY 501 TTGATTTTGAAGCCCGGATATATGCGCAATGCAAGCGAAGCAAGCCACTTTCAC 560
DB 3376 TCAGGTTTGTGATGCTGGATATTTGGCAATGTAGAGATCCAGCTGACCTTACCC 3435
QY 561 ATTCAAGCCAGAGCCCGGATGATGCTGCAATCAACACCTTCTGAGGCTTACCTCC 620
DB 3436 ACAGCTGCAACTAGAGCACTTATGCTTGAGAGGAACACCTTCTGAGGCTTACCTCC 3495
QY 621 GCATCCGATTCGATCTTGGAGAGCAATGCAAGCTTCTTCAACACATCAACCTTCC 680
DB 3496 GAATAGATTGAAATTTTGGAGGCAATCCAGGCTTTCAGCAACATCAATCTCCAT 3555
QY 681 GTTACCCCTCCGTGGAGCAATCCAGGCTTCCGCTGACCTTTCCTCATATACATCG 740
DB 3556 GTTTCGCACTAGAGGCTGCAATGACACATCTGTATCTTGTACTTGTCTCATATG 3615
QY 741 CCGGCTCTTACCGGCGCGGCGGCAATGCAAGCGGCTTGGCGGCGGAGAGCTCTCA 800
DB 3616 CCGGTTTCTTAACTGATAGGCGCAACTGCAAGGCTTGTGCTGCTGCGAGCATTTCTGA 3675
QY 801 ATGCGGAAAAGGCTTCTTGTAGTGGAGTTGAAAGTGGATTTCTGAGATTAACGCA 860
DB 3676 ATGCTAAGGAAGCTTGTGAATTTGGCCCAACATTTGCTTCTGACTTCTTCACTTCAACCTGA 3735
QY 861 AAGAAGGCTAGCACTTGTAAAGCGGCGGCGGCTTGGCGGCGGAGAGCTCTCACTTC 920
DB 3736 AAGAAGGCTAGCACTTGTAAAGCGGCGGCGGCTTGGCGGCGGAGAGCTCTCACTTC 920
QY 921 TATTGATCTAATGCTTGTGCTTGTGCGAGCTTGTAGGCGGCTTGTGCTTGTGCTG 3795
DB 3796 TCTTTGAAGGAACATCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 3855
QY 981 TTATGCAAGGGAAGCGGAGCTTATGCGATCTTACACACAAATTTGAAGCATCAACCTG 1040
DB 3856 TGATGCAAGGGAAGCGGAGCTTATGCGATCTTACACACAAATTTGAAGCATCAACCTG 1040
QY 1041 GTCAAATGAGGCGGCGGCGGCGGCTTATGCGATCTTACACACAAATTTGAAGCATCA 1100
DB 3916 GCGAGATGAGGCTGCTGCTATATATGAGACATTTTGGATGGAAGCTTATATCAAG 3975
QY 1101 CGGCGGAAAAGGCTTACAGAAATGAGCGGCTTACAGAAACCAAGACATGCTTATGCTC 1160
DB 3976 CTCTTAAGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4035
QY 1161 TCGGATATCTCCGATATGCTGAACTGAACTGAACTGAACTGAACTGAACTGAACTG 1220
DB 4036 TTGAACTTACCCGCAATGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 4095
QY 1221 TGAATCAGAGGGAATTCATCCGTCAGACGACACCAATGATGAGCTTTCAGAAACA 1280
DB 4096 CAATTTAGAGGAGATCACTCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 4155
QY 1281 AAGCTTACAGGCTGATCTTCAAGCAACCCCAATGCAATGCAATGCAATGCAATGCAATG 1340
DB 4156 AGGCTTACAGGCTGATCTTCAAGCAACCCCAATGCAATGCAATGCAATGCAATGCA 4215
QY 1341 GTCTGCGCATTTGTAATGGAATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1400
DB 4216 GTTGGCTATTTGCTCAATTTGAAACTCATGTTTGTGCTCAATTTGCTGCTGCTGCTG 4275
QY 1401 ATTTCACAAAGATTTACATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1460
DB 4276 ACTATATCAATATGAGTGTGCTTATCTCATGCGAGAGAAACCCAGCTTGAAT 4335
QY 1461 ACGGCTCAAAAGCTGCAAAATGCGCATGCTTCTTACTGCTTCTGAGCTTCTGCTGCTG 1520
DB 1520

DB 4336 ATGCTTTCAAGGAGCTGCAAAATTTGCATGCGATCTTACTTCTGTAACCTCAATATTTG 4395
QY 1521 CAATTCAGATTCACATCTTATGCTTCAAGGCGGAGCAACCAATCAAGATTTATTTCTC 1580
DB 4396 CTAACTCCGATTAAGGAGCTATGCTCAAAAGGCGGAGCAACCAATCAAGATTTATTTCTC 4455
QY 1581 TCGGATTAATTTGACCGAGCAAAACCCGAGAGAGCTGCAATCTTAAACTGATGCTG 1640
DB 4456 TGGGATTTATTTATCTATGAGAAACCAATGAGGCTTGTGATGCTTCACTTATTTGCTT 4515
QY 1641 GCACTTATTTGCTGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTT 1700
DB 4516 CACTTTCTCTGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATG 4575
QY 1701 AATGACAGTGAAGCAACCCGATGAGGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1760
DB 4576 AGAACCTGTAAGAAATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4635
QY 1761 ACGGAGCTGCGACCCGCTGAGATTTGCGAGAAAGATCTGCTGCTGCTGCTGCTGCTG 1820
DB 4636 ATGAGAGATCTCCACCTTCAAGATTTGTGAGAAAGCTGCTGCTGCTGCTGCTGCTG 4695
QY 1821 AATAGCTTCTGCTTATGAGAGAGCTTTCAGGAGGAGATTCAGGATTTATGAGAGAG 1880
DB 4696 AGTACATCTTCTGCTTATGAGAGAGCTTTCAGGAGGAGATTCAGGATTTATGAGAGAG 4755
QY 1881 TCCGACAGCTTCTGCTTATGAGAGAGCTTTCAGGAGGAGATTCAGGATTTATGAGAGAG 1940
DB 4756 TAAAGGAGCTTCTGCTTATGAGAGAGCTTTCAGGAGGAGATTCAGGATTTATGAGAGAG 4815
QY 1941 CCTCAGCTTCTGCTTATGAGAGAGCTTTCAGGAGGAGATTCAGGATTTATGAGAGAG 2000
DB 4816 GCTTATCTTCTGCTTATGAGAGAGCTTTCAGGAGGAGATTCAGGATTTATGAGAGAG 4875
QY 2001 AAGTTAAGCTTATGAGAGAGCTTTCAGGAGGAGATTCAGGATTTATGAGAGAG 2060
DB 4876 AAGTTAAGCTTATGAGAGAGCTTTCAGGAGGAGATTCAGGATTTATGAGAGAG 4935
QY 2061 AAGTTAAGCTTATGAGAGAGCTTTCAGGAGGAGATTCAGGATTTATGAGAGAG 2120
DB 4936 AAGTTAAGCTTATGAGAGAGCTTTCAGGAGGAGATTCAGGATTTATGAGAGAG 4995
QY 2121 TGAACGAGAAAGGCTGAGCTGCGGAGAGAGAGCTTTCAGGAGGAGATTCAGGAGAG 2180
DB 4996 TGAACGAGAAAGGCTGAGCTGCGGAGAGAGAGCTTTCAGGAGGAGATTCAGGAGAG 5055
QY 2181 GCAAAAGCTTATGAGAGAGCTTTCAGGAGGAGATTCAGGAGGAGATTCAGGAGAG 2240
DB 5056 GCAAAAGCTTATGAGAGAGCTTTCAGGAGGAGATTCAGGAGGAGATTCAGGAGAG 5115
QY 2241 TTGCAATATGCTAG 2254
DB 5116 TTGCAATATGCTAG 5129

RESULT 9
AAC61194
ID AAC61194 standard: DNA: 4415 BP.
AC AAC61194;
XX
XX 12-FEB-2001 (first entry)
DE Pea PEPAL1 DNA sequence SEQ ID 9.
XX
XX
KW Pea: promoter; plant: prevention; pathogen: infection; ds.
OS Pisum sativum.
XX
XX JP2000245463-A.
XX
XX
XX
XX PD 12-SRP-2000.
XX

OY 1518 TCGCAATCCAGTCCACCAACATGTTGAAAGCGCCGACCAACACATGACAGCTTAAT 1577
 DB 2126 TAGTAATATCTAGTACCAAGATATGTAAGATATGATATGATATGATATGATATGATAT 2185
 OY 1578 CTCTCGCATTAATTTCCAGGAGAGAAACCCGACAGCATGTCATCTTAAACATGAT 1637
 DB 2186 CTCTCGCATTAATTTCCAGGAGAGAAACCCGACAGCATGTCATCTTAAACATGAT 2245
 OY 1638 GGTGCAATATCTAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1697
 DB 2246 CTCTCGCATTAATTTCCAGGAGAGAAACCCGACAGCATGTCATCTTAAACATGAT 2305
 OY 1698 TGAATTCAGACAGTGAAGACACCGTAAAGCAGCATGTCATCTTAAACATGAT 1757
 DB 2306 TCAAGAAACAGTCTCAAGAAATCTGATGATGATGATGATGATGATGATGATGATGAT 2365
 OY 1758 TCAAGGCGACCTCCACCGCTGAGATGATGATGATGATGATGATGATGATGATGATGAT 1817
 DB 2366 ACAAGAAACAGTCTCAAGAAATCTGATGATGATGATGATGATGATGATGATGATGATGAT 2425
 OY 1818 GTCATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1877
 DB 2426 GAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2485
 OY 1878 AGCTCCAGACGCTTCTGTCGACACACCGCTTAACACAGCGGACGACGACGACGACGAC 1937
 DB 2486 AACTTAAGAGGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2545
 OY 1938 ACACGTCACATCTCCAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1997
 DB 2546 GTTTGTTGTTTGTGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2605
 OY 1998 AGAGGTTGAAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2057
 DB 2606 AGGAGTGAAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2665
 OY 2058 TTTAAAGCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2117
 DB 2666 TCAAGGAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2725
 OY 2118 TTTTGAACCGGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2177
 DB 2726 TGTCTACCGGCGGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2785
 OY 2178 TGTGCAAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2237
 DB 2786 TGTGCAAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2845
 OY 2238 CTCTCGCATTAATTTCCAGGAGAGAAACCCGACAGCATGTCATCTTAAACATGAT 2297
 DB 2846 CCATCCCATTAATTTCCAGGAGAGAAACCCGACAGCATGTCATCTTAAACATGAT 2905
 OY 2298 GTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 2331
 DB 2906 CTCTCGCATTAATTTCCAGGAGAGAAACCCGACAGCATGTCATCTTAAACATGAT 2939

RESULT 12

AAO33019 standard; cDNA; 1707 BP.

XX AAO33019;
 AC 07-MAY-1993 (first entry)
 XX Pisum sativum L. PAL gene.
 XX Phenylalanine ammonia lyase; phenyl propanoid isochlorogenic acid; plant;
 KM pea; ss.
 XX Pisum sativum.
 OS

PN JP04330285-A.
 PD 18-NOV-1992.
 XX 26-APR-1991; 91JP-0097697.
 XX 26-APR-1991; 91JP-0097697.
 XX (TAKS) TAKASAGO PERFUMERY CO LTD.
 DR WPI: 1993-003497/01.
 XX New phenylalanine ammonia lyase gene from *Pisum sativum* L.
 PT plasmid contg. gene, and *E. coli* transformed with plasmid, for
 PT prepn. of lyase.
 XX Disclousure: Page 8; 9pp; Japanese.
 PS The sequence is that of the phenylalanine ammonia lyase (PAL) gene
 CC from *Pisum sativum* L. The gene may be used to produce PAL,
 CC recombinantly. Plants contg. the PAL gene can contain a large amt.
 CC of phenyl propanoid isochlorogenic acid.
 CC See also AAO31985.
 XX
 SQ Sequence 1707 BP; 496 A; 345 C; 377 G; 489 T; 0 other;

Query Match 34.4%; Score 841; DB 14; Length 1707;
 Best Local Similarity 72.6%; Pred. No. 1.5e-11;
 Matches 1101; Conservative 0; Mismatches 415; Indels 1; Gaps 1;

OY 737 ATGCGCGGCTCTTAACCG 795
 DB 1 ATGCGCGGCTCTTAACCG 800
 OY 796 CCGTCAATGCGGAAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 855
 DB 61 ACTTACTGCTAAGCAAGCTTTGAAATGCGCTAACATCAATTTGATGATGATGATGATGATGAT 910
 OY 856 GCGGAAGAAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 915
 DB 121 ACCCAAGAAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 970
 OY 916 CGTCTATTTGATGCTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 975
 DB 181 AGTCTCTTGTATGCAAAATATAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1030
 OY 976 TGAAGTATGCAAGGAAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1035
 DB 241 CCAAGTATGCAAGGAAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1090
 OY 1036 CCGTGTCAATGCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1095
 DB 301 CCGTGTCAATGCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1150
 OY 1096 CAAAGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1155
 DB 361 GAAAGTGTGTAAGAAAGCTGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1210
 OY 1156 TGTCTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1215
 DB 421 TGTCTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1270
 OY 1216 CAAATGATGCAAGGAAAGCTGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1275
 DB 481 CAAATGATGCAAGGAAAGCTGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1330
 OY 1276 AACAAGGCTTACACGCGGTAATCTGCAAGGAAAGCGGCGGCGGCGGCGGCGGCGGCGGCG 1335
 DB 541 GAAACAAAGGCTTACACGCGGTAATCTGCAAGGAAAGCGGCGGCGGCGGCGGCGGCGGCGG 1390
 OY 1336 CACCGGCTGCGCATGCTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 1395
 DB 601 CACCGGCTGCGCATGCTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 1450

Db 66 TTCTCTGAGGCTGTTTCTGAGGCTATGAAATGAGAGGATTTTCTGAGGATATCTA 125
 Oy 1017 CACACAAATTCAGGATTCACCCCTGCTCAATTCGAGGCGCGCGGATCATGAGATATTTT 1076
 Db 126 CACACAACTCAGGATTCACCCCTGCTCAATTCGAGGCGCGCGGATCATGAGATATTTT 185
 Oy 1077 TGGAGGAGGATTTTCTGAGGCTATGAAATGAGAGGATTTTCTGAGGATATCTA 1136
 Db 186 TTGAAGGAGGATTTTCTGAGGCTATGAAATGAGAGGATTTTCTGAGGATATCTA 245
 Oy 1137 AACGAAACAGATTCGATTCGCTCTCTGATATCTGCTGATATCTGCTGATATCTG 1196
 Db 246 AGCTAAAGATATGATATCTGATATCTGATATCTGATATCTGATATCTGATATCTG 305
 Oy 1197 AAGTATTCGATTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1256
 Db 306 AGGTTATTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 365
 Oy 1257 GATTGATTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1316
 Db 366 CAGTATTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 425
 Oy 1317 TCGGAGGATTCGATTCGCTCTCTGATATCTGCTGATATCTGCTGATATCTGCTG 1376
 Db 426 TGGAGGAGGATTTTCTGAGGCTATGAAATGAGAGGATTTTCTGAGGATATCTA 485
 Oy 1377 CCGAATTCGATTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1436
 Db 486 CCGAATTCGATTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 545
 Oy 1437 GTGAGGATTCGATTCGCTCTCTGATATCTGCTGATATCTGCTGATATCTGCTG 1496
 Db 546 GTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 605
 Oy 1497 ACTGTTTCGATTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1556
 Db 606 ACTGTTTCGATTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 665
 Oy 1557 AACGAAACAGATTCGATTCGCTCTCTGATATCTGCTGATATCTGCTGATATCTGCTG 1616
 Db 666 AACGAAACAGATTCGATTCGCTCTCTGATATCTGCTGATATCTGCTGATATCTGCTG 725
 Oy 1617 TGGAGGAGGATTTTCTGAGGCTATGAAATGAGAGGATTTTCTGAGGATATCTA 1676
 Db 726 TGGAGGAGGATTTTCTGAGGCTATGAAATGAGAGGATTTTCTGAGGATATCTA 785
 Oy 1677 TACGAGGATTCGATTCGCTCTCTGATATCTGCTGATATCTGCTGATATCTGCTG 1736
 Db 786 TACGAGGATTCGATTCGCTCTCTGATATCTGCTGATATCTGCTGATATCTGCTG 845
 Oy 1737 AAAAGGATTCGATTCGCTCTCTGATATCTGCTGATATCTGCTGATATCTGCTG 1796
 Db 846 GGAAGGATTCGATTCGCTCTCTGATATCTGCTGATATCTGCTGATATCTGCTG 905
 Oy 1797 ATCTGCTCTGATTCGATTCGCTCTCTGATATCTGCTGATATCTGCTGATATCTGCTG 1856
 Db 906 ACTCTGCTCTGATTCGATTCGCTCTCTGATATCTGCTGATATCTGCTGATATCTGCTG 965
 Oy 1857 GATCTGCTCTGATTCGATTCGCTCTCTGATATCTGCTGATATCTGCTGATATCTGCTG 1916
 Db 966 CAGTATTCGATTCGATTCGCTCTCTGATATCTGCTGATATCTGCTGATATCTGCTG 1025
 Oy 1917 GCGAAGGATTCGATTCGCTCTCTGATATCTGCTGATATCTGCTGATATCTGCTG 1976
 Db 1026 GCGAAGGATTCGATTCGCTCTCTGATATCTGCTGATATCTGCTGATATCTGCTG 1085
 Oy 1977 AATGAGGATTCGATTCGCTCTCTGATATCTGCTGATATCTGCTGATATCTGCTG 2036
 Db 1086 AATGAGGATTCGATTCGCTCTCTGATATCTGCTGATATCTGCTGATATCTGCTG 1145
 Oy 2037 CAGTATTCGATTCGATTCGCTCTCTGATATCTGCTGATATCTGCTGATATCTGCTG 2096
 Db 1146 CAGTATTCGATTCGATTCGCTCTCTGATATCTGCTGATATCTGCTGATATCTGCTG 1205

Oy 2097 GAGGAGGATTCGATTCGCTCTCTGATATCTGCTGATATCTGCTGATATCTGCTG 2156
 Db 1206 GAGGAGGATTCGATTCGCTCTCTGATATCTGCTGATATCTGCTGATATCTGCTG 1265
 Oy 2157 TGGAGGAGGATTTTCTGAGGCTATGAAATGAGAGGATTTTCTGAGGATATCTA 2216
 Db 1266 TGGAGGAGGATTTTCTGAGGCTATGAAATGAGAGGATTTTCTGAGGATATCTA 1325
 Oy 2217 TGGAGGAGGATTTTCTGAGGCTATGAAATGAGAGGATTTTCTGAGGATATCTA 2276
 Db 1326 TGGAGGAGGATTTTCTGAGGCTATGAAATGAGAGGATTTTCTGAGGATATCTA 1385
 Oy 2277 GATGATTCGATTCGATTCGCTCTCTGATATCTGCTGATATCTGCTGATATCTGCTG 2332
 Db 1386 AAGGAGGATTCGATTCGCTCTCTGATATCTGCTGATATCTGCTGATATCTGCTG 1441

RESULT 15
 ID AAA68154
 AAA68154 standard; DNA; 1520 bp.
 AC AAA68154;
 XX
 LI 24-OCT-2000 (first entry)
 XX
 DE Pinus radiata PAL nucleotide sequence SEQ ID NO:247.
 XX
 KM Plant, ligin, ligin biosynthetic pathway, Eucalyptus grandis;
 KM Pinus radiata; Monterey pine; ds.
 XX
 OS Pinus radiata.
 XX
 PN WO200022099-A1.
 XX
 XX 20-APR-2000.
 PD
 XX 06-OCT-1999; 99WO-NZ00168.
 PF
 PR 09-OCT-1998; 98US-0169789.
 PR 14-JUL-1999; 99US-0143811.
 XX
 PA (GENE-) GENESIS RES & DEV CORP LTD.
 PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
 XX
 PI Bloksberg LN, Havukkala IJ;
 DR
 WP: 2000-317962/27
 PS
 XX
 CC The present invention describes isolated polynucleotides and proteins
 CC encoding and representing the enzymes cinnamate 4-hydroxylase (C4H),
 CC cinnamate 3-hydroxylase (C3H), phenylase (PNI), O-methyl transferase
 CC (OMT), cinnamyl alcohol dehydrogenase (CAD), cinnamoyl-CoA reductase
 CC (CCR), phenylalanine ammonia-lyase (PAL), 4-cinnamate-CoA ligase (4CL),
 CC coniferol glucosyl transferase (CGT), coniferin beta-glucosidase (CBG),
 CC laccase, peroxidase, ferulate-5-hydroxylase (F5H), alpha-amylase,
 CC caffeic acid methyl transferase, caffeoyl CoA methyl transferase,
 CC cinnamate CoA ligase, cytochrome P450 3XX1A, diphenyl oxylase, flavanol
 CC glucosyl transferase, flavanol hydroxylase, and isolavone reductase,
 CC which are involved in the ligin biosynthetic pathway. The
 CC polynucleotides can be used for mediating ligin content, ligin
 CC composition and the structure of a plant, especially eucalyptus and pine
 CC species, and for modifying the activity of an enzyme involved in ligin
 CC biosynthetic pathway, and for producing a plant having altered ligin
 CC content, composition and structure. They can be used for designing probes
 CC and primers useful for detecting similar DNA and RNA sequences in any

organism and for PCR amplification. The lignin content can be efficiently modified using the polynucleotides. AA67908 to AA68201 and AB16341 to CC AA16449 represent the polynucleotide and protein sequences used in the CC exemplification of the present invention.

Sequence 1520 BP; 389 A; 356 C; 376 G; 399 T; 0 other:

Query Match 26.3%; Score 642; DB 21; Length 1520;

Best local similarity 67.2%; Pred. No. 6-2e-106;

Matches 939; Conservative 0; Mismatches 455; Indels 4; Gaps 2;

QY 714 GTACCTTGTCCATATACATACATGACCTTAAACGACCCCAACTCCAAAG 773
 DB 1 GTATCTGGTTCCTCCCTCTTATATGTCTGGCTCTTACCGGAGGCTTATTCACAG 60
 QY 774 CCGTGGCCCGCCAGAGAGTCATGCGCAAAAGGCTTCTGCTGACCGCGAGTTG 833
 DB 61 TCAGATCCAGAGATGGAAATGGAATGAGCGGAGCGGAGCGCTCAAGAAATGGGCTGG 120
 QY 834 AAGGTGGTCTTGGAGTTACAGCGCGAAGAGGCTAGCACTTGTAAAGGACCGCGG 893
 DB 121 AA---AAGCCCTTGAATTTGCAACCTTAAGAGAGGTCTGCCATATGCAATGGCACTTCAG 177
 QY 894 TGGAGTCCGAGATGCTTCATGCTCTATTTGATGCTATCTACTGCTGTGTGCG 953
 DB 178 TGGAGAGCACTGGCTCCATGTTGTGTTGATGCGCAATGTTCTGCTCTCTGTG 237
 QY 954 AAGGTATGCGGATCTTGGAGTTCAGAGTATGCAAGGAGCGGAGTGTAAAGGACCT 1013
 DB 238 AAGTATCTGCTGATCTTGTGAGTATATATGATGAGAGGTTAGATTACAGATCAT 297
 QY 1014 TGACACATTAATTTGAATATATATATATATATATATATATATATATATATAT 1073
 DB 298 TAACCTCAAG 357
 QY 1074 TTTGGAGGAG 1133
 DB 358 TCTTGGAGGAGAG-TCTTATATGAAACAGCTGCTAGCTCATAGATGAAATCTCTGCG 416
 QY 1134 AGAAGCAAAACAG 1193
 DB 417 AGAAGCAAAAG 476
 QY 1194 TCGAGTATATGAT 1253
 DB 477 TCGAGATATATGAT 536
 QY 1254 ACCGATGATGAG 1313
 DB 537 ATCCAGATATGAT 596
 QY 1314 CAATCGAGTTCATATGAG 1373
 DB 597 CTAATGCTGTTCCATGAG 656
 QY 1374 TCGCTAATTTTGTGAG 1433
 DB 657 TCGCTAATTTTGTGAG 716
 QY 1434 CCGGTGAG 1493
 DB 717 GTGTGAG 776
 QY 1494 CTTAGCTTCTGAGCTTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1553
 DB 777 CTTAGCTTCTGAGCTTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 836
 QY 1554 AACAGCAATATAG 1613
 DB 837 AACAGCAATATAG 896
 QY 1614 CAGTGCATCTTAAATCATATATATATATATATATATATATATATATATATATAT 1673

DB 897 CATTGATATATCTGAAGCTATGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 956
 QY 1674 ATTATGAGCATTTGGAG 1733
 DB 957 ATTATGAGCATTTGGAG 1016
 QY 1734 CTAAG 1793
 DB 1017 CTAAG 1076
 QY 1794 AAGATCTCTCCGCTGTTGTTGATCGTGAATACGCTTCGCTTACATGAGAGAGAGAG 1853
 DB 1077 AAGATCTCTCCGCTGTTGTTGATCGTGAATACGCTTCGCTTACATGAGAGAGAGAG 1136
 QY 1854 GCGGACATACCATTAATGAG 1913
 DB 1137 ATGCCAGCTATCCCATTTGACTCAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1196
 QY 1914 AGGAG 1973
 DB 1197 AGGAG 1256
 QY 1974 AAGATTTGAAG 2033
 DB 1257 CCGAGCTGAAG 1316
 QY 2034 ATGATTTGATGATTTCCAAACAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2093
 DB 1317 GAG 1376
 QY 2094 TANGAG 2111
 DB 1377 TGAAG 1394

Search completed: March 28, 2003, 21:56:09
 Job time : 556 secs

GenCode version 5.1 A_J5_4574
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protocol - protein search, using SW model

Run on: March 28, 2003, 23:41:56, Search time 67 seconds

(without alignments)
1414.048 Million cell updates/sec

Title: US-09-964-992a-1
Perfect score: 3644
Sequence: 1 MENCNHNVCVNEIKIKPL.....IIIPPLEIIOGDMGSGPLAPIC 711

Scoring table:
Gapop 10 0, Gapext 0 5
BLOSUM62

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: A-Geneset_101002.*
1: /SID52/qcqdala/genescq/genescq-emb1/AA1981.DAT.*
2: /SID52/qcqdala/genescq/genescq-emb1/AA1981.DAT.*
3: /SID52/qcqdala/genescq/genescq-emb1/AA1982.DAT.*
4: /SID52/qcqdala/genescq/genescq-emb1/AA1983.DAT.*
5: /SID52/qcqdala/genescq/genescq-emb1/AA1984.DAT.*
6: /SID52/qcqdala/genescq/genescq-emb1/AA1985.DAT.*
7: /SID52/qcqdala/genescq/genescq-emb1/AA1986.DAT.*
8: /SID52/qcqdala/genescq/genescq-emb1/AA1987.DAT.*
9: /SID52/qcqdala/genescq/genescq-emb1/AA1988.DAT.*
10: /SID52/qcqdala/genescq/genescq-emb1/AA1989.DAT.*
11: /SID52/qcqdala/genescq/genescq-emb1/AA1990.DAT.*
12: /SID52/qcqdala/genescq/genescq-emb1/AA1991.DAT.*
13: /SID52/qcqdala/genescq/genescq-emb1/AA1992.DAT.*
14: /SID52/qcqdala/genescq/genescq-emb1/AA1993.DAT.*
15: /SID52/qcqdala/genescq/genescq-emb1/AA1994.DAT.*
16: /SID52/qcqdala/genescq/genescq-emb1/AA1995.DAT.*
17: /SID52/qcqdala/genescq/genescq-emb1/AA1996.DAT.*
18: /SID52/qcqdala/genescq/genescq-emb1/AA1997.DAT.*
19: /SID52/qcqdala/genescq/genescq-emb1/AA1998.DAT.*
20: /SID52/qcqdala/genescq/genescq-emb1/AA1999.DAT.*
21: /SID52/qcqdala/genescq/genescq-emb1/AA2000.DAT.*
22: /SID52/qcqdala/genescq/genescq-emb1/AA2001.DAT.*
23: /SID52/qcqdala/genescq/genescq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	3644	100.0	711 23 AA097130	Lettuce LSPAL1 enz
2	3294	90.4	712 23 AA097131	Lettuce LSPAL2 enz
3	3059.5	84.0	777 14 AAR30077	Pistum sativum L.P
4	3045.5	83.6	725 21 AAG36709	Arabidopsis thalia
5	3044.5	83.5	709 21 AAG36710	Arabidopsis thalia
6	3041.5	83.5	698 21 AAG36711	Arabidopsis thalia
7	3004	82.4	711 21 AAG54103	Arabidopsis thalia
8	3004	82.4	717 21 AAG54102	Arabidopsis thalia
9	2951	81.0	681 21 AAG54104	Arabidopsis thalia
10	2907.5	79.8	736 21 AAG53399	Arabidopsis thalia

11	2807	77.0	699 21 AAG53400	Arabidopsis thalia
12	2648	72.7	648 21 AAG53401	Arabidopsis thalia
13	1892	51.9	423 21 AAG51947	Arabidopsis thalia
14	1799	49.4	424 21 AAB15402	Eucalyptus grandis
15	1779	48.8	495 23 ABB78982	Lolium perenne Lp
16	1752	48.1	396 21 AAG21948	Arabidopsis thalia
17	1709	46.9	386 21 AAG21949	Arabidopsis thalia
18	1475	40.5	489 23 ABB78983	Lolium perenne Lp
19	1345	36.9	412 21 AAB16404	Pinus radiata PAL
20	1337	36.7	304 23 AA012056	Protein encoded by R. glutinis PAL po
21	1025	28.1	720 23 ABB07690	R. glutinis phenyl l-phenylalanine am
22	946.5	27.3	716 23 AAE29564	Sequence of l-phen
23	944.5	27.3	716 9 AAP83149	Sequence of Rhodosp
24	944.5	27.3	716 9 AAP83149	Rhodotorula glutin
25	944.5	27.3	716 23 AAE20653	R. toruloides PAL
26	944.5	27.3	716 23 ABB07643	Rhodotorula glutin
27	944.5	27.3	716 23 AAB16384	l-phenylalanine am
28	944.5	27.3	716 10 AAE95783	R. glutinis phenyl
29	932.5	27.2	716 23 AAE20655	Rhodotorula glutin
30	932.5	27.2	716 23 AAE20654	Rhodotorula glutin
31	931.5	27.2	716 23 AAE20659	Rhodotorula glutin
32	931.5	27.2	716 23 AAE20657	R. glutinis phenyl
33	931.5	27.2	716 23 AAE20658	Rhodotorula glutin
34	931.5	27.2	716 23 AAE20658	Rhodotorula glutin
35	931.5	27.2	716 23 AAE20658	Rhodotorula glutin
36	930.5	27.2	716 23 AAE20658	Rhodotorula glutin
37	930.5	27.2	716 23 AAE20658	Rhodotorula glutin
38	988	27.1	686 23 AAE20657	Rhodotorula glutin
39	987.5	27.1	716 16 AAB70682	Stabilised phenyl
40	987.5	27.1	716 23 AAE20656	R. glutinis phenyl
41	946.5	27.0	716 23 AAE20653	Rhodotorula glutin
42	946.5	27.0	716 23 AAE20653	Rhodotorula glutin
43	984.5	26.9	716 23 AAE20711	R. glutinis phenyl
44	940.5	26.9	716 23 AAE20588	Rhodotorula glutin
45	971.5	26.7	713 23 ABB07692	R. rubra/mucillagin

ALIGNMENTS

RESULT 1
AA097130
ID AA097130 standard; Protein: 711 AA.
AC AA097130;
XX
XX 13-APR-2002 (first entry)
PT
XX
XX lettuce LSPAL1 enzyme.
DE
XX
XX lettuce LSPAL2 enzyme.
KW
XX lettuce; phenylalanine ammonia-lyase; plant wounding; plant damage;
KW plant disease; plant browning; vegetable crop; plant; LSPAL1;
XX enzyme.
XX
XX
XX lettuce sativa.
OS
XX
XX W0200226928-A2.
XX
XX
XX 04-APR-2002.
XX
XX 26-SEP-2001; 2001WO US10192.
XX
XX 26-SEP-2001; 2000US-235956P.
XX
XX (REGC) UNIV CALIFORNIA.
XX
XX Salveit M, Campos R, Norogaki H, Suslow T;
PI WPI: 2002-435152/46.
DR N-PSDB: ABR50679.
XX
XX Novel polynucleotide useful in preparing plant expression constructs to
PT modify the response of a plant cell to wounding, damage or other injury

from disease causing organisms or from plant pests -
 Claim 1: Fig 6; 46pp; English.

The present invention relates to the isolation of novel lettuce (Lactuca sativa) phenylalanine ammonia-lyase (LsPAL) enzymes, and the polynucleotide sequences encoding them. An antibody that binds to a PAL enzyme is useful for measuring the relative amount of PAL levels in a tissue. The polynucleotide sequences encoding PAL are useful in the preparation of plant expression constructs for modifying features of the response of a plant cell to wounding, damage or other injury from disease. Disruption or down-regulation of PAL activity can be used to reduce the browning response to wounding in vegetable crops. The present sequence represents lettuce LsPAL1 enzyme.

Sequence 711 AA;

Query Match 100.0%; Score 3644; DB 23; Length 711;
 Best local similarity 100.0%; Pred. No. 2.3e-283;
 Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MENCNHNAGVYVNELCIKDPLNMGVAAPALTGSILDEYKKMAVEFRKRYVVKLGGFTLTVSQ 60
 DB 1 MENCNHNAGVYVNELCIKDPLNMGVAAPALTGSILDEYKKMAVEFRKRYVVKLGGFTLTVSQ 60
 QY 61 VAGIAANDSDTVKVELSEARAGVYKASDWMESNNKGTDSYGVTTGFGATSHRTKQG 120
 DB 61 VAGIAANDSDTVKVELSEARAGVYKASDWMESNNKGTDSYGVTTGFGATSHRTKQG 120
 QY 121 GALOKELIRFLNAGIFGNGTETSHLPHSATRAMIVRINTLGGYSGIRFELLTAITKF 180
 DB 121 GALOKELIRFLNAGIFGNGTETSHLPHSATRAMIVRINTLGGYSGIRFELLTAITKF 180
 QY 181 LNNNTTPTCLPRGTTTASGDLVPLSYTAGLLTGRNSKAVPTGCEVINAERKAFAGVGS 240
 DB 181 LNNNTTPTCLPRGTTTASGDLVPLSYTAGLLTGRNSKAVPTGCEVINAERKAFAGVGS 240
 QY 241 GFPELOPEKREGIALVNGTAVSGMAVLFEDANVALLSEVLSAIFAEVMOGKPEFTDHLT 300
 DB 241 GFPELOPEKREGIALVNGTAVSGMAVLFEDANVALLSEVLSAIFAEVMOGKPEFTDHLT 300
 QY 301 HKLKHHPQOIEAAMIMEYILDGSYVKAQKVHEMDPLQKDRVALRTSPOMLTPOLE 360
 DB 301 HKLKHHPQOIEAAMIMEYILDGSYVKAQKVHEMDPLQKDRVALRTSPOMLTPOLE 360
 QY 361 VIRSTTKMIEKEINSVNDNPIIDVSRKKALHGNFQSTPIGVSMNDTRLAIAIGKLMFA 420
 DB 361 VIRSTTKMIEKEINSVNDNPIIDVSRKKALHGNFQSTPIGVSMNDTRLAIAIGKLMFA 420
 QY 421 QFSELVNDYFVNNGLPSNLISGRNPSLDYGFKGGEIAMAAYSCSELOFLANVTNHHVQSAQ 480
 DB 421 QFSELVNDYFVNNGLPSNLISGRNPSLDYGFKGGEIAMAAYSCSELOFLANVTNHHVQSAQ 480
 QY 481 HNOVYNSIGLISARKTAEAVDILKMSSTYLVALCOSTLDRHEENKSTVKNVTSQVAK 540
 DB 481 HNOVYNSIGLISARKTAEAVDILKMSSTYLVALCOSTLDRHEENKSTVKNVTSQVAK 540
 QY 541 KVLTMGVNCELHPRCFCKTLIPVYREYVPAVILHWCSSTYPLMOKIKQVLVDAALNG 600
 DB 541 KVLTMGVNCELHPRCFCKTLIPVYREYVPAVILHWCSSTYPLMOKIKQVLVDAALNG 600
 QY 601 ETEKNTNISIFOKIATFEELKVLIPKEVEGVRIAYENDLSIPNBKACRSYPLREVR 660
 DB 601 ETEKNTNISIFOKIATFEELKVLIPKEVEGVRIAYENDLSIPNBKACRSYPLREVR 660
 QY 661 EELGRGFLTGKVTSPGEFDFPVFTAMCKQIILPPLLELTAGVNGEPLPIC 711
 DB 661 EELGRGFLTGKVTSPGEFDFPVFTAMCKQIILPPLLELTAGVNGEPLPIC 711

RESUJ.T 2
 AA097131
 ID AA097131 standard; Protein: 712 AA.

XX AA097131;
 AC 13-AUG-2002 (first entry)
 XX
 DT
 XX
 DE Lettuce LsPAL2 enzyme.
 XX
 KW Lettuce; phenylalanine ammonia-lyase; plant wounding; plant damage;
 XX plant disease; plant browning, vegetable crop, plant; LsPAL2,
 XX enzyme.
 OS Lactuca sativa.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 21..22
 FT /note= "Encoded by CUGAAGCG"

XX W0200226028-A2.
 XX
 XX 04-APR-2002.
 XX
 XX 26-SEP-2001; 2001WO-US30192.
 XX
 XX 26-SEP-2000; 2000US-235956P.
 XX
 XX (REGC) UNIV CALIFORNIA.
 XX
 XX Saltvelt M, Campos R, Norugaki H, Suslow T;
 XX
 XX WPI: 2002-435152/46.
 XX
 XX N-PSDB: ABR50680.

PT Novel polynucleotide useful in preparing plant expression constructs to
 modify the response of a plant cell to wounding, damage or other injury
 from disease causing organisms or from plant pests -

Claim 2: Fig 8; 46pp; English.

The present invention relates to the isolation of novel lettuce (Lactuca sativa) phenylalanine ammonia-lyase (LsPAL) enzymes, and the polynucleotide sequences encoding them. An antibody that binds to a PAL enzyme is useful for measuring the relative amount of PAL levels in a tissue. The polynucleotide sequences encoding PAL are useful in the preparation of plant expression constructs for modifying features of the response of a plant cell to wounding, damage or other injury from disease. Disruption or down-regulation of PAL activity can be used to reduce the browning response to wounding in vegetable crops. The present sequence represents lettuce LsPAL2 enzyme.

Sequence 712 AA;

Query Match 90.4%; Score 3294; DB 23; Length 712;
 Best local similarity 90.7%; Pred. No. 2.8e-255;
 Matches 644; Conservative 29; Mismatches 33; Indels 4; Gaps 2;

QY 1 MENCNHNAGVYVNELCIKDPLNMGVAAPALTGSILDEYKKMAVEFRKRYVVKLGGFTLTVSQ 60
 DB 6 MENCNHNAGVYVNELCIKDPLNMGVAAPALTGSILDEYKKMAVEFRKRYVVKLGGFTLTVSQ 65
 QY 61 VAGIAANDSDTVKVELSEARAGVYKASDWMESNNKGTDSYGVTTGFGATSHRTKQG 120
 DB 61 VAGIAANDSDTVKVELSEARAGVYKASDWMESNNKGTDSYGVTTGFGATSHRTKQG 124
 QY 121 GALOKELIRFLNAGIFGNGTETSHLPHSATRAMIVRINTLGGYSGIRFELLTAITKF 180
 DB 121 GALOKELIRFLNAGIFGNGTETSHLPHSATRAMIVRINTLGGYSGIRFELLTAITKF 184
 QY 181 LNNNTTPTCLPRGTTTASGDLVPLSYTAGLLTGRNSKAVPTGCEVINAERKAFAGVGS 240
 DB 181 LNNNTTPTCLPRGTTTASGDLVPLSYTAGLLTGRNSKAVPTGCEVINAERKAFAGVGS 244
 QY 241 GFPELOPEKREGIALVNGTAVSGMAVLFEDANVALLSEVLSAIFAEVMOGKPEFTDHLT 300
 DB 241 GFPELOPEKREGIALVNGTAVSGMAVLFEDANVALLSEVLSAIFAEVMOGKPEFTDHLT 300

DB 245 GFEFLQPKHGLAVNGTAVGSGMASMVLPIANVIALLSFVLSALFAVWQCKPPTDHLT 304
 QY 301 IRLKLIHPGYLEAAALMEYILLGSDYVKAAGVHEKUPYLQKKQDQYALRTSPQWLGQTE 360
 DB 305 IRLKLIHPGYLEAAALMEYILLGSDYVKAAGVHEKUPYLQKKQDQYALRTSPQWLGQTE 364
 QY 361 VLRSTKMLEPINSVINNPILIVSRKAIHCGNPGCPPTGVMNINRIALAAIGKLMFA 420
 DB 365 VLRSTKMLEPINSVINNPILIVSRKAIHCGNPGCPPTGVMNINRIALAAIGKLMFA 424
 QY 421 QFSELYNDFYNNGLPSNLGSEFSLDYGFKAQFEIAMAASVSELYFLANVINIIVSAGU 480
 DB 425 QFSELYNDFYNNGLPSNLGSEFSLDYGFKAQFEIAMAASVSELYFLANVINIIVSAGU 484
 QY 481 IHDQVNSLGLSARKFAVAVIOLKMSSTYLVALVQSTIDPLEFPNNKSTYKKNVSVYAK 540
 DB 485 IHDQVNSLGLSARKFAVAVIOLKMSSTYLVALVQSTIDPLEFPNNKSTYKKNVSVYAK 544
 QY 541 KVLTMGVNGLHPSRFCKKDLKVDREYVFAVYIDVCSCTYPLMOKIKQYVVDHALNNG 600
 DB 545 KLITTSVNGELHPSRCEKDLRVVDREYFAVYIDVCSCTYPLMOKIKQYVVDHALNNG 603
 QY 601 FTEKNTNISFOKIAFPEBELKVLIPKEVSPVIAVENDTLSTPNPKAPSPYPLFPVR 660
 DB 604 --FENDAGTSPFKISPEPEELKAVLIPKEVSPVIAVENDTLSTPNPKAPSPYPLFPVR 661
 QY 661 FELGKRFELTGEKVTSPGFEPDPVPTAMKQIOLPDLPLCTLGMMNCEPLPT 710
 DB 662 FELGKRFELTGEKVTSPGFEPDPVPTAMKQIOLPDLPLCTLGMMNCEPLPT 711

RESULT 3
 AAR30077
 ID AAR30077 standard: Protein: 777 AA.
 XX
 AC AAR30077:
 DT 07-MAY-1993 (first entry)
 XX
 DE Pisum sativum L. PAL.
 XX
 KM phenylalanine ammonia lyase; phenyl propionid isoflavonoid; plant;
 KW pea.
 XX
 OS Pisum sativum.
 XX
 PN JP04330285-A.
 XX
 PD 18-NOV-1992.
 XX
 PE 26-APR-1991: 91JP-0097697.
 XX
 PR 26-APR-1991: 91JP-0097697
 XX
 PA (TAKS) TAKASAGO PHARMACY CO LTD
 XX
 DR WPI: 1993-003497/01.
 DR N-PSDB: AAQ31985.
 XX
 DT New phenylalanine ammonia lyase gene from Pisum sativum L.
 DT plasmid contg. gene, and E. coli transformed with plasmid, for
 PT prep. of lysate.
 XX
 PS Disclousure: Page 7; 9pp; Japanese.
 XX
 CC The sequence is that of phenylalanine ammonia lyase (PAL)
 CC from Pisum sativum L.. The gene encoding PAL may be used to
 CC produce PAL recombinantly. Plants contg. the PAL gene can
 CC contain a large amt. of phenyl propanoid isoflavonoids.
 XX
 SQ Sequence 777 AA:
 Query Match RA 0%: Score 1059.5; DB 14; Length 777;

Host Local Similarity 83.8%; Prod. No. 2,1e-216;
 Matches 594, Conservative 57; Mismatches 51; Indels 7; Gaps 4;
 QY 3 NNNHNVGVNDELCKDPLNMAVAEALTSLLDEKRVKVAEHEKPVKLGQELLIVNQA 62
 DB 74 NAKNNMMKVN---ADPLNMGVAAAEAMKGSHTDEKRVKVAEHEKPVKLGQELLIVNQA 140
 QY 63 GAAANDSITVYVFLSPAPACVAKASSIDVMSNNKKGIDSGVITGCAISHPETKQGA 122
 DB 131 AIAAIDHG--YVVELSESPAPACVAKASSIDVMSNNKKGIDSGVITGCAISHPETKQGA 188
 QY 123 LQKELIRFLNAGIPGNGTETSHLPKSAIRAAAIIVRINTLQGYSGIRFLFLAIFKFN 182
 DB 189 LQKELIRFLNAGIPGNGTETSHLPKSAIRAAAIIVRINTLQGYSGIRFLFLAIFKFN 248
 QY 183 NNITPCLPLRGITTAAGLVPLSTAGLLTTPPSKAVAGPTGEVLAKEKFAAGVGGF 242
 DB 249 NNVTPL--LKGITPASCGRVPLSYLAGITGPSPSKAGHSGPILNAKFAAGVAGF 407
 QY 243 FFIQPKHGLAVNGTAVGSGMASMVLPIANVIALLSFVLSALFAVWQCKPPTDHLTHK 302
 DB 308 FFIQPKHGLAVNGTAVGSGMASMVLPIANVIALLSFVLSALFAVWQCKPPTDHLTHK 367
 QY 303 LKHHPQULIAAAIMEYILLGSDYVKAAGVHEKUPYLQKKQDQYALRTSPQWLGQTE 362
 DB 368 LKHHPQULIAAAIMEYILLGSDYVKAAGVHEKUPYLQKKQDQYALRTSPQWLGQTE 427
 QY 363 BSTKMKIEPINSVNDNPLIVSRKAIHCGNPGCPPTGVMNINRIALAAIGKLMFA 422
 DB 428 BSTKMKIEPINSVNDNPLIVSRKAIHCGNPGCPPTGVMNINRIALAAIGKLMFA 487
 QY 423 SELVNDYFNNGLPSNLGSEFSLDYGFKAQFEIAMAASVSELYFLANVINIIVSAGU 481
 DB 488 SELVNDYFNNGLPSNLGSEFSLDYGFKAQFEIAMAASVSELYFLANVINIIVSAGU 547
 QY 482 NQIVNSGLISAPKTAFAVAVIKIMSSIVYVAGQSTPHPEFTNPKSTYFNINIVSAGU 541
 DB 548 NQIVNSGLISAPKTAFAVAVIKIMSSIVYVAGQSTPHPEFTNPKSTYFNINIVSAGU 607
 QY 542 VITMGVNEFLHPSRCEKDLRVVDREYFAVYIDVCSCTYPLMOKIKQYVVDHALNNG 601
 DB 608 VITMGVNEFLHPSRCEKDLRVVDREYFAVYIDVCSCTYPLMOKIKQYVVDHALNNG 667
 QY 602 TREKNTNISFOKIAFPEBELKVLIPKEVSPVIAVENDTLSTPNPKAPSPYPLFPVR 661
 DB 668 TREKNTNISFOKIAFPEBELKVLIPKEVSPVIAVENDTLSTPNPKAPSPYPLFPVR 727
 QY 662 ELGRGFLTGKAVTSPGEPHDPVPTAMKQIOLPDLPLCTLGMMNCEPLPT 710
 DB 728 ELGRGFLTGKAVTSPGEPHDPVPTAMKQIOLPDLPLCTLGMMNCEPLPT 776

RESULT 4
 AAG36709
 ID AAG36709 standard: Protein: 725 AA.
 XX
 AC AAG36709:
 DT 18-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SHU 10 No: 45026.
 XX
 KM Protein identification: signal transduction pathway; metabolic pathway;
 KM hybridisation assay; genetic mapping; gene expression control; promoter;
 KM termination sequence.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PR 25-FEB-2000; 2000HP-0301439.

XX
PR 25-FEB-1999; 9905-0121825-
PR 05-MAR-1999; 9905-0123180-
PR 09-MAR-1999; 9905-0123548-
PR 23-MAR-1999; 9905-0125788-
PR 25-MAR-1999; 9905-0126264-
PR 29-MAR-1999; 9905-0126785-
PR 01-APR-1999; 9905-0127462-
PR 06-APR-1999; 9905-0128234-
PR 08-APR-1999; 9905-0128714-
PR 16-APR-1999; 9905-0129845-
PR 19-APR-1999; 9905-0130077-
PR 21-APR-1999; 9905-0130449-
PR 23-APR-1999; 9905-0130510-
PR 23-APR-1999; 9905-0130891-
PR 28-APR-1999; 9905-0131449-
PR 30-APR-1999; 9905-0132048-
PR 30-APR-1999; 9905-0132407-
PR 04-MAY-1999; 9905-0132484-
PR 05-MAY-1999; 9905-0132485-
PR 06-MAY-1999; 9905-0132486-
PR 07-MAY-1999; 9905-0132487-
PR 11-MAY-1999; 9905-0132863-
PR 14-MAY-1999; 9905-0134256-
PR 14-MAY-1999; 9905-0134218-
PR 14-MAY-1999; 9905-0134219-
PR 14-MAY-1999; 9905-0134221-
PR 14-MAY-1999; 9905-0134370-
PR 18-MAY-1999; 9905-0134768-
PR 19-MAY-1999; 9905-0134941-
PR 20-MAY-1999; 9905-0135124-
PR 21-MAY-1999; 9905-0135353-
PR 24-MAY-1999; 9905-0135629-
PR 25-MAY-1999; 9905-0136021-
PR 27-MAY-1999; 9905-0136392-
PR 28-MAY-1999; 9905-0136782-
PR 01-JUN-1999; 9905-0137222-
PR 03-JUN-1999; 9905-0137528-
PR 04-JUN-1999; 9905-0137502-
PR 07-JUN-1999; 9905-0137724-
PR 08-JUN-1999; 9905-0138094-
PR 10-JUN-1999; 9905-0138540-
PR 10-JUN-1999; 9905-0138847-
PR 14-JUN-1999; 9905-0139119-
PR 16-JUN-1999; 9905-0139452-
PR 16-JUN-1999; 9905-0139453-
PR 17-JUN-1999; 9905-0139492-
PR 18-JUN-1999; 9905-0139454-
PR 18-JUN-1999; 9905-0139455-
PR 18-JUN-1999; 9905-0139456-
PR 18-JUN-1999; 9905-0139457-
PR 18-JUN-1999; 9905-0139458-
PR 18-JUN-1999; 9905-0139459-
PR 18-JUN-1999; 9905-0139460-
PR 18-JUN-1999; 9905-0139461-
PR 18-JUN-1999; 9905-0139462-
PR 18-JUN-1999; 9905-0139463-
PR 18-JUN-1999; 9905-0139750-
PR 18-JUN-1999; 9905-0139763-
PR 21-JUN-1999; 9905-0139817-
PR 22-JUN-1999; 9905-0139899-
PR 23-JUN-1999; 9905-0140353-
PR 23-JUN-1999; 9905-0140354-
PR 24-JUN-1999; 9905-0140695-
PR 28-JUN-1999; 9905-0140823-
PR 29-JUN-1999; 9905-0140991-
PR 30-JUN-1999; 9905-0141287-
PR 01-JUL-1999; 9905-0141842-
PR 01-JUL-1999; 9905-0142154-
PR 02-JUL-1999; 9905-0142055-
PR 06-JUL-1999; 9905-0142390-
PR 08-JUL-1999; 9905-0142803-
PR 09-JUL-1999; 9905-0142920-

PR 12-JUL-1999; 9905-0142977-
PR 13-JUL-1999; 9905-0143542-
PR 14-JUL-1999; 9905-0143624-
PR 15-JUL-1999; 9905-0144005-
PR 16-JUL-1999; 9905-0144085-
PR 16-JUL-1999; 9905-0144086-
PR 19-JUL-1999; 9905-0144325-
PR 19-JUL-1999; 9905-0144331-
PR 19-JUL-1999; 9905-0144332-
PR 19-JUL-1999; 9905-0144333-
PR 19-JUL-1999; 9905-0144334-
PR 19-JUL-1999; 9905-0144335-
PR 20-JUL-1999; 9905-0144352-
PR 20-JUL-1999; 9905-0144632-
PR 21-JUL-1999; 9905-0144814-
PR 21-JUL-1999; 9905-0144814-
PR 21-JUL-1999; 9905-0145086-
PR 21-JUL-1999; 9905-0145088-
PR 22-JUL-1999; 9905-0145085-
PR 22-JUL-1999; 9905-0145087-
PR 22-JUL-1999; 9905-0145089-
PR 23-JUL-1999; 9905-0145192-
PR 23-JUL-1999; 9905-0145145-
PR 23-JUL-1999; 9905-0145218-
PR 23-JUL-1999; 9905-0145224-
PR 26-JUL-1999; 9905-0145276-
PR 27-JUL-1999; 9905-0145913-
PR 27-JUL-1999; 9905-0145918-
PR 27-JUL-1999; 9905-0145919-
PR 28-JUL-1999; 9905-0145951-
PR 02-AUG-1999; 9905-0146386-
PR 02-AUG-1999; 9905-0146388-
PR 02-AUG-1999; 9905-0146389-
PR 02-AUG-1999; 9905-0147038-
PR 04-AUG-1999; 9905-0147204-
PR 04-AUG-1999; 9905-0147302-
PR 05-AUG-1999; 9905-0147192-
PR 05-AUG-1999; 9905-0147260-
PR 06-AUG-1999; 9905-0147303-
PR 06-AUG-1999; 9905-0147416-
PR 09-AUG-1999; 9905-0147493-
PR 09-AUG-1999; 9905-0147935-
PR 10-AUG-1999; 9905-0148171-
PR 11-AUG-1999; 9905-0148319-
PR 12-AUG-1999; 9905-0148341-
PR 13-AUG-1999; 9905-0148341-
PR 13-AUG-1999; 9905-0148685-
PR 16-AUG-1999; 9905-0148684-
PR 17-AUG-1999; 9905-0149175-
PR 18-AUG-1999; 9905-0149426-
PR 20-AUG-1999; 9905-0149722-
PR 20-AUG-1999; 9905-0149723-
PR 20-AUG-1999; 9905-0149929-
PR 23-AUG-1999; 9905-0149902-
PR 23-AUG-1999; 9905-0149930-
PR 25-AUG-1999; 9905-0150566-
PR 26-AUG-1999; 9905-0150884-
PR 27-AUG-1999; 9905-0151065-
PR 27-AUG-1999; 9905-0151066-
PR 27-AUG-1999; 9905-0151080-
PR 30-AUG-1999; 9905-0151303-
PR 31-AUG-1999; 9905-0151438-
PR 01-SEP-1999; 9905-0151930-
PR 07-SEP-1999; 9905-0152363-
PR 10-SEP-1999; 9905-0153070-
PR 13-SEP-1999; 9905-0153758-
PR 15-SEP-1999; 9905-0154018-
PR 16-SEP-1999; 9905-0154039-
PR 20-SEP-1999; 9905-0154779-
PR 22-SEP-1999; 9905-0155139-
PR 23-SEP-1999; 9905-0155486-
PR 24-SEP-1999; 9905-0155659-
PR 28-SEP-1999; 9905-0156458-

```

PR 29-SEP-1999: 9905-0156596.
PR 04-OCT-1999: 9905-0157117.
PR 05-OCT-1999: 9905-0157753.
PR 06-OCT-1999: 9905-0157865.
PR 07-OCT-1999: 9905-0158029.
PR 08-OCT-1999: 9905-0158232.
PR 12-OCT-1999: 9905-0158369.
PR 13-OCT-1999: 9905-0159294.
PR 13-OCT-1999: 9905-0159294.
PR 13-OCT-1999: 9905-0159294.
PR 14-OCT-1999: 9905-0159329.
PR 14-OCT-1999: 9905-0159330.
PR 14-OCT-1999: 9905-0159331.
PR 14-OCT-1999: 9905-0159637.
PR 14-OCT-1999: 9905-0159638.
PR 18-OCT-1999: 9905-0159584.
PR 21-OCT-1999: 9905-0160741.
PR 21-OCT-1999: 9905-0160747.
PR 21-OCT-1999: 9905-0160766.
PR 21-OCT-1999: 9905-0160770.
PR 21-OCT-1999: 9905-0160814.
PR 21-OCT-1999: 9905-0160815.
PR 22-OCT-1999: 9905-0160980.
PR 22-OCT-1999: 9905-0160981.
PR 22-OCT-1999: 9905-0160989.
PR 25-OCT-1999: 9905-0161404.
PR 25-OCT-1999: 9905-0161405.
PR 25-OCT-1999: 9905-0161406.
PR 26-OCT-1999: 9905-0161359.
PR 26-OCT-1999: 9905-0161360.
PR 26-OCT-1999: 9905-0161361.
PR 28-OCT-1999: 9905-0161920.
PR 28-OCT-1999: 9905-0161992.
PR 28-OCT-1999: 9905-0161993.
PR 29-OCT-1999: 9905-0162142.

```

Query Match 83.6%; Score 3045.5; DB 21; Length 725;
 Best Local Similarity 81.4%; Pred. No. 2,562,35;
 Matches 589, Conservative 52, Mismatches 66, Indels 17, Gaps 3;

```

QY 3 NGNHVN---GVNRLC-----IKDPLNMGVAAFALITQSHIDRYKKMAEPRKP 47
DB 4 NGAHKSNGSGVAMLCGIDIKTKNNVINAEIPLNMFAAEQMKSHIDEVPMVAEPRKP 63
QY 48 VVAKIGETITVSQVACIAAANISDITKVELSTAAKAGYKASSDWMVSMNKGUSYGVTF 107
DB 64 VVNLGGETITIGQVAAISITIGNS--VKVPLSTAPAGYNASSDWMVSMNKGUSYGVTF 121
QY 108 GRTATSHPTKQYALDYELIFLNAIFGNTTETSHITLPHSATRAMIVPINTLLQYS 167
DB 122 GRTATSHPTKQYALDYELIFLNAIFGNTTETSHITLPHSATRAMIVPINTLLQYS 181
QY 168 GTRFELIFATFKLNNNTTPQLRGTTTASQDLYPLSYIAGLITGRNSKAVGPTGVYL 227
DB 182 GTRFELIFATFKLNNNTTPQLRGTTTASQDLYPLSYIAGLITGRNSKAVGPTGVYL 241
QY 228 NAKKAAAGVHGFELQPRKQALVNGTAVGSCMASMVLPEANVLAISEVLSATIPAE 287
DB 242 TAFKAKIAGIASSGPEPDQPRKQALVNGTAVGSCMASMVLPEANVLAISEVLSATIPAE 301
QY 288 VMGKPEETDHLTIKLKHPRQGLEAATIMEYILDOGSDVYKAAQVHENDPLQKQJRYA 347
DB 302 VMGKPEETDHLTIKLKHPRQGLEAATIMEYILDOGSDVYKAAQVHENDPLQKQJRYA 361
QY 348 LRTSPQMLAQHOIVLRSSSTKMEIREINSVNDNPLIDVSRNKALQGNFOGTPIGVSMINT 407
DB 362 LRTSPQMLAQHOIVLRSSSTKMEIREINSVNDNPLIDVSRNKALQGNFOGTPIGVSMINT 421
QY 408 KLAIAAIGKLMIAQFSEIYNDYVNGGLISNLSGGKRNISLDYQFGKGLAMASYSSEIQLT 467
DB 422 KLAIAAIGKLMIAQFSEIYNDYVNGGLISNLSGGKRNISLDYQFGKGLAMASYSSEIQLT 481
QY 468 ANPVTNHYGSAEQHNDVNSIGLISARKEAVNDILKLMSSYIYVALQSIDLHREEM 527

```

```

DB 482 ANPVTNHYGSAEQHNDVNSIGLISARKEAVNDILKLMSSYIYVALQSIDLHREEM 541
QY 528 KSTVKNVSOVAKKAVITGNGVCELPSPKCEKDLKAVDREYVAVIDVCSQIYPLMOK 587
DB 542 KQTVKNVSOVAKKAVITGNGVCELPSPKCEKDLKAVDREYVAVIDVCSQIYPLMOK 601
QY 588 LQVLYVLAALNGEETEKNTNISFOKIAFEEELKVLIPKEVEGVIAVENDLISINRT 647
DB 602 LQVLYVLAALNGEETEKNTNISFOKIAFEEELKVLIPKEVEGVIAVENDLISINRT 661
QY 648 KQPSYPIYPRFPFELGSGPTTGEKVTSGDEPDVFPFAMKQSIDPLECTGAWNGEP 707
DB 662 KQPSYPIYPRFPFELGSGPTTGEKVTSGDEPDVFPFAMKQSIDPLECTGAWNGEP 721
QY 708 LPIG 711
DB 722 LPIG 725
RESULT 5
ID AAG36710 standard; protein; 709 AA.
AC AAG36710;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 45027.
XX
KW hybridisation; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999: 9905-0121825.
XX 05-MAR-1999: 9905-0123180.
XX 09-MAR-1999: 9905-0123548.
XX 23-MAR-1999: 9905-0125788.
XX 25-MAR-1999: 9905-0126264.
XX 29-MAR-1999: 9905-0126785.
XX 01-APR-1999: 9905-0127462.
XX 06-APR-1999: 9905-0128234.
XX 08-APR-1999: 9905-0128714.
XX 16-APR-1999: 9905-0129845.
XX 19-APR-1999: 9905-0130077.
XX 21-APR-1999: 9905-0130449.
XX 23-APR-1999: 9905-0130610.
XX 23-APR-1999: 9905-0130891.
XX 28-APR-1999: 9905-0131449.
XX 30-APR-1999: 9905-0132048.
XX 04-MAY-1999: 9905-0132407.
XX 05-MAY-1999: 9905-0132484.
XX 06-MAY-1999: 9905-0132485.
XX 06-MAY-1999: 9905-0132486.
XX 07-MAY-1999: 9905-0132487.
XX 11-MAY-1999: 9905-0134256.
XX 14-MAY-1999: 9905-0134219.
XX 14-MAY-1999: 9905-0134221.
XX 14-MAY-1999: 9905-0134370.
XX 18-MAY-1999: 9905-0134768.
XX 19-MAY-1999: 9905-0134941.
XX 20-MAY-1999: 9905-0135124.
XX 21-MAY-1999: 9905-0135353.

```

[illegible]

FR	02-AUG-1999	9905-0146369
FR	03-AUG-1999	9905-0147038
FR	04-AUG-1999	9905-0147302
FR	04-AUG-1999	9905-0147302
FR	05-AUG-1999	9905-0147192
FR	05-AUG-1999	9905-0147260
FR	05-AUG-1999	9905-0147303
FR	05-AUG-1999	9905-0147416
FR	09-AUG-1999	9905-0147849
FR	09-AUG-1999	9905-0147935
FR	10-AUG-1999	9905-0148131
FR	11-AUG-1999	9905-0148319
FR	12-AUG-1999	9905-0148341
FR	13-AUG-1999	9905-0148684
FR	16-AUG-1999	9905-0149176
FR	17-AUG-1999	9905-0149175
FR	18-AUG-1999	9905-0149426
FR	20-AUG-1999	9905-0149742
FR	20-AUG-1999	9905-0149723
FR	20-AUG-1999	9905-0149829
FR	21-AUG-1999	9905-0149902
FR	23-AUG-1999	9905-0149936
FR	25-AUG-1999	9905-0150560
FR	26-AUG-1999	9905-0150884
FR	27-AUG-1999	9905-0151065
FR	27-AUG-1999	9905-0151060
FR	30-AUG-1999	9905-0151303
FR	31-AUG-1999	9905-0151348
FR	01-SEP-1999	9905-0151930
FR	07-SEP-1999	9905-0152363
FR	10-SEP-1999	9905-0153079
FR	13-SEP-1999	9905-0153758
FR	15-SEP-1999	9905-0154018
FR	16-SEP-1999	9905-0154039
FR	20-SEP-1999	9905-0154779
FR	22-SEP-1999	9905-0155131
FR	24-SEP-1999	9905-0155484
FR	24-SEP-1999	9905-0155569
FR	28-SEP-1999	9905-0156458
FR	29-SEP-1999	9905-0156596
FR	04-OCT-1999	9905-0157117
FR	05-OCT-1999	9905-0157785
FR	06-OCT-1999	9905-0157865
FR	07-OCT-1999	9905-0158632
FR	08-OCT-1999	9905-0158832
FR	12-OCT-1999	9905-0159293
FR	13-OCT-1999	9905-0159293
FR	13-OCT-1999	9905-0159294
FR	14-OCT-1999	9905-0159331
FR	14-OCT-1999	9905-0159533
FR	14-OCT-1999	9905-0159637
FR	14-OCT-1999	9905-0159638
FR	18-OCT-1999	9905-0159884
FR	21-OCT-1999	9905-0160741
FR	21-OCT-1999	9905-0160768
FR	21-OCT-1999	9905-0160770
FR	21-OCT-1999	9905-0160815
FR	22-OCT-1999	9905-0160880
FR	22-OCT-1999	9905-0160981
FR	25-OCT-1999	9905-0160989
FR	26-OCT-1999	9905-0161405
FR	26-OCT-1999	9905-0161405
FR	26-OCT-1999	9905-0161406
FR	26-OCT-1999	9905-0161360
FR	26-OCT-1999	9905-0161361
FR	26-OCT-1999	9905-0161360

PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145132.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147408.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.

PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155446.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159239.
PR 13-OCT-1999; 99US-0159244.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159654.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 83.58; Score 3041.5; DB 21; Length 698;

Best Local Similarity 81.28; Fred. No. 4.9e-235;

Matches 584; Conservative 52; Mismatches 61; Indels 5; Caps 2;

QY 10 VNELCICKPLNMGVAAEALTGSHLDEVKKMAEERKPVYKLGGETLTVSOVAGIAAND 69
DB 2 VIN--AEDPLNMGAAAEQMGKSHLDEVKKMAEERKPVYKLGGETLTVSOVAGIAAND 58
QY 70 SDYKVELSEARACVAKVASSDVMESNMKCTDSYGVITGCGATSHKRTKQCCALQKELIR 129
DB 59 S-YKVELSEARACVAKVASSDVMESNMKCTDSYGVITGCGATSHKRTKQCCALQKELIR 116
QY 130 PLNMGIFGNGTSTHTLPHSATRAAMIYRIINTLLOGSGIRFELTALTKFLNNNTPL 189
DB 117 PLNMGIFGNGTSTHTLPHSATRAAMIYRIINTLLOGSGIRFELTALTKFLNNNTPL 176
QY 190 PLRGITFASGDLVPLSYLAGLITGRPNKAVGPTGEVLAENKAFAAAGVGGEFFELQPRE 249
DB 177 PLRGITFASGDLVPLSYLAGLITGRPNKAVGPTGEVLAENKAFAAAGVGGEFFELQPRE 236
QY 250 GLIALVNGIAVSSGMAWVLFUANYLALLSVLSAIFAEVQCKPEFIDILIRKLAHPQV 309
DB 237 GLIALVNGIAVSSGMAWVLFUANYLALLSVLSAIFAEVQCKPEFIDILIRKLAHPQV 296
QY 310 IEAAAMEYIIDGSDYVKAOKVHEMDLQPKODRYALKTSQWIGPOIEVIRSTTKI 369
DB 297 IEAAAMEYIIDGSDYVKAOKVHEMDLQPKODRYALKTSQWIGPOIEVIRSTTKI 356
QY 370 EPEINSVNDPLDIVSRNKALHGGNFGTPIGVSMNTPLAIAICKLMAFQFSELVNDF 429
DB 357 EPEINSVNDPLDIVSRNKALHGGNFGTPIGVSMNTPLAIAICKLMAFQFSELVNDF 416

XX HP1033405-A2.
PN
XX
PD 06-SEP-2000.
XX
PP 25-FEB-2000; 2000EP-0301434
XX
PR 25-FEB-1999; 9905-0121825.
PR 05-MAR-1999; 9905-0123180.
PR 09-MAR-1999; 9905-0123548.
PR 23-MAR-1999; 9905-0125788.
PR 25-MAR-1999; 9905-0126264.
PR 29-MAR-1999; 9905-0126785.
PR 01-APR-1999; 9905-0127462.
PR 06-APR-1999; 9905-0128234.
PR 08-APR-1999; 9905-0128714.
PR 16-APR-1999; 9905-0129845.
PR 19-APR-1999; 9905-0130077.
PR 21-APR-1999; 9905-0130449.
PR 23-APR-1999; 9905-0130510.
PR 23-APR-1999; 9905-0130891.
PR 28-APR-1999; 9905-0131449.
PR 30-APR-1999; 9905-0132048.
PR 30-APR-1999; 9905-0132407.
PR 04-MAY-1999; 9905-0132484.
PR 05-MAY-1999; 9905-0132485.
PR 06-MAY-1999; 9905-0132486.
PR 06-MAY-1999; 9905-0132487.
PR 07-MAY-1999; 9905-0132863.
PR 11-MAY-1999; 9905-0134256.
PR 14-MAY-1999; 9905-0134218.
PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134221.
PR 14-MAY-1999; 9905-0134376.
PR 18-MAY-1999; 9905-0134768.
PR 19-MAY-1999; 9905-0134941.
PR 20-MAY-1999; 9905-0135124.
PR 21-MAY-1999; 9905-0135353.
PR 24-MAY-1999; 9905-0135629.
PR 25-MAY-1999; 9905-0136021.
PR 27-MAY-1999; 9905-0136392.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137222.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 10-JUN-1999; 9905-0138847.
PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139492.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139464.
PR 18-JUN-1999; 9905-0139465.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139899.
PR 23-JUN-1999; 9905-0140353.
PR 23-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140645.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140991.
PR 30-JUN-1999; 9905-0141287.

PR 01-JUL-1999; 9905-0141842.
PR 01-JUL-1999; 9905-0142154.
PR 02-JUL-1999; 9905-0142055.
PR 06-JUL-1999; 9905-0142390.
PR 08-JUL-1999; 9905-0142803.
PR 09-JUL-1999; 9905-0142920.
PR 12-JUL-1999; 9905-0142977.
PR 13-JUL-1999; 9905-0143542.
PR 14-JUL-1999; 9905-0143624.
PR 15-JUL-1999; 9905-0144005.
PR 16-JUL-1999; 9905-0144086.
PR 19-JUL-1999; 9905-0144325.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144632.
PR 20-JUL-1999; 9905-0144684.
PR 21-JUL-1999; 9905-0144814.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145088.
PR 22-JUL-1999; 9905-0145085.
PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145192.
PR 23-JUL-1999; 9905-0145145.
PR 23-JUL-1999; 9905-0145218.
PR 23-JUL-1999; 9905-0145224.
PR 26-JUL-1999; 9905-0145276.
PR 27-JUL-1999; 9905-0145913.
PR 27-JUL-1999; 9905-0145918.
PR 27-JUL-1999; 9905-0145919.
PR 28-JUL-1999; 9905-0145951.
PR 02-AUG-1999; 9905-0146386.
PR 02-AUG-1999; 9905-0146389.
PR 02-AUG-1999; 9905-0146389.
PR 03-AUG-1999; 9905-0147038.
PR 04-AUG-1999; 9905-0147204.
PR 04-AUG-1999; 9905-0147302.
PR 05-AUG-1999; 9905-0147192.
PR 05-AUG-1999; 9905-0147260.
PR 06-AUG-1999; 9905-0147303.
PR 06-AUG-1999; 9905-0147416.
PR 09-AUG-1999; 9905-0147493.
PR 09-AUG-1999; 9905-0147935.
PR 10-AUG-1999; 9905-0148171.
PR 11-AUG-1999; 9905-0148319.
PR 12-AUG-1999; 9905-0148341.
PR 13-AUG-1999; 9905-0148565.
PR 13-AUG-1999; 9905-0148684.
PR 16-AUG-1999; 9905-0149368.
PR 17-AUG-1999; 9905-0149175.
PR 18-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149722.
PR 20-AUG-1999; 9905-0149723.
PR 20-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149902.
PR 23-AUG-1999; 9905-0149930.
PR 25-AUG-1999; 9905-0150560.
PR 26-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151066.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151438.
PR 01-SEP-1999; 9905-0151930.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.

```

PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157785.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161982.
PR 28-OCT-1999; 99US-0161983.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 82.48; Score 3004; Len 71; Length 717;
 Best local similarity 81.58; pred. No. 5.2e-232;
 Matches 577; Conservative 52; Mismatches 77; Indels 2; Gaps 1;

```

QY 4 GNHGVVNNELCTKDPINMGVAAEALTGSHLDEYKKKVAEERKRYKLGSETLTVSQAAG 63
DB 12 GKTAVAVTTTKTLADPLINMGIAADOMKSHLDEYKKKVAEERKRYKLGSETLTVSQAAG 71
QY 64 IAAANDSTTVKVELSEAFARVAVKSSDMWESNMKGTDSYVTTTTPATSHPTKQJUAL 123
DB 72 ISTVGS--KVELAETSRVAVKSSDMWESNMKGTDSYVTTTTPATSHPTKQJUAL 129
QY 124 QKEILKFTNAGIPGNCIEISHILPHSAIKRAMIKINILAGYSGILFEELIAIKFPLN 183
DB 130 QTELIFPLNATIPINTKETCTITPQSATRAAMLVAVNTLLQYSGIFELLEITSLIHH 189
QY 184 NITPLPLPRTITASGDLVPLSTIAGLLTPPNKSKAVPTGEVULNAKKAFAAAGVEGGEF 243
DB 190 NISPSILPKGTITFASGDLVPLSTIAGLLTPPNKSKAVPTGEVULNAKKAFAAAGVEGGEF 249
QY 244 ELOPEEGIALVNGTAVGSMASVLPEDANVLLISEVLSITFEVNGQKPEFDHLTIHL 303
DB 250 DLQPEEGIALVNGTAVGSMASVLPEDANVLLISEVLSITFEVNGQKPEFDHLTIHL 309
QY 304 KHHQQLIEAALIMEYILDGSDYVKAACKVHBMPLQKPKQDKAKIKISPVMLQSLQLEVL 363
DB 310 KHHQQLIEAALIMEYILDGSDYVKAACKVHBMPLQKPKQDKAKIKISPVMLQSLQLEVL 369
QY 364 SSTKMEIEINSVNDPLIDVSRKKALHGNFQGTPIGVSMONTRLAIAIGKLMFAQFS 423
DB 370 QATKSIEIEINSVNDPLIDVSRKKALHGNFQGTPIGVSMONTRLAIAIAIGKLMFAQFS 429

```

```

QY 424 ELVNDPYNGLPSNLSCGRNPSLDYGFKGGEIAMASYCSHLOPLAMPVNNHVSQAHOHQ 483
DB 430 ELVNDPYNGLPSNLSTASSNESLDYGFKGGEIAMASYCSHLOPLAMPVNNHVSQAHOHQ 489
QY 484 DVNSICGLSARKTAEAVIDILKMSSTYVALCOSIDLRLHLENNKSTVNTVSQAARKVL 543
DB 490 DVNSICGLSARKTAEAVIDILKMSSTYVALCOSIDLRLHLENNKSTVNTVSQAARKVL 549
QY 544 TMAYNAGELHPSKFEKELLVVDPEYFAYIDVNSSTYPLMKLVYVMAINKGETE 603
DB 550 TTGINGELHPSKFEKELLVVDPEYFAYIDVNSSTYPLMKLVYVMAINKGETE 609
QY 604 KNTNTSIFOKIATFEEDLKVLLPEKEGVRJAVENDTSLIPNTKACRSYPLRYVREEL 663
DB 610 KNAVTSIFOKIATFEEDLKVLLPEKEGVRJAVENDTSLIPNTKACRSYPLRYVREEL 669
QY 664 GRGELTGKVTSPGEHEDRYVTAMCKGQIITDPLIFGLCGMNGRPLDTC 711
DB 670 GIKLITGKVTSPGEHEDRYVTAMCKGQIITDPLIFGLCGMNGRPLDTC 717

```

RESULT 9

AAG54104
 ID AAG54104 standard; Protein; 681 AA.

AC AAG54104;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment S6010 No. 68950.

KW Protein identification: signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.

OS Arabidopsis thaliana.

FN EPI033405-A2.

XX EPI033405-A2.

XX 06 SEP-2000.

```

XX 25-FEB-2000; 2000RP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.

```

PR 20-MAY-1999: 99US-0135124
PR 21-MAY-1999: 99US-0135353
PR 24-MAY-1999: 99US-0135629
PR 25-MAY-1999: 99US-0136021
PR 27-MAY-1999: 99US-0136392
PR 28-MAY-1999: 99US-0136782
PR 01-JUN-1999: 99US-0137222
PR 03-JUN-1999: 99US-0137526
PR 04-JUN-1999: 99US-0137502
PR 07-JUN-1999: 99US-0137724
PR 08-JUN-1999: 99US-0138094
PR 10-JUN-1999: 99US-0138540
PR 10-JUN-1999: 99US-0138847
PR 14-JUN-1999: 99US-0139119
PR 16-JUN-1999: 99US-0139452
PR 16-JUN-1999: 99US-0139453
PR 17-JUN-1999: 99US-0139492
PR 18-JUN-1999: 99US-0139454
PR 18-JUN-1999: 99US-0139455
PR 18-JUN-1999: 99US-0139456
PR 18-JUN-1999: 99US-0139457
PR 18-JUN-1999: 99US-0139458
PR 18-JUN-1999: 99US-0139459
PR 18-JUN-1999: 99US-0139460
PR 18-JUN-1999: 99US-0139461
PR 18-JUN-1999: 99US-0139462
PR 18-JUN-1999: 99US-0139463
PR 18-JUN-1999: 99US-0139750
PR 18-JUN-1999: 99US-0139763
PR 21-JUN-1999: 99US-0139817
PR 22-JUN-1999: 99US-0139899
PR 23-JUN-1999: 99US-0140353
PR 23-JUN-1999: 99US-0140354
PR 24-JUN-1999: 99US-0140695
PR 28-JUN-1999: 99US-0140823
PR 28-JUN-1999: 99US-0140991
PR 30-JUN-1999: 99US-0141287
PR 01-JUL-1999: 99US-0141842
PR 01-JUL-1999: 99US-0142154
PR 02-JUL-1999: 99US-0142055
PR 06-JUL-1999: 99US-0142390
PR 08-JUL-1999: 99US-0142803
PR 09-JUL-1999: 99US-0142920
PR 12-JUL-1999: 99US-0142977
PR 13-JUL-1999: 99US-0143542
PR 14-JUL-1999: 99US-0143624
PR 15-JUL-1999: 99US-0144005
PR 16-JUL-1999: 99US-0144085
PR 16-JUL-1999: 99US-0144086
PR 19-JUL-1999: 99US-0144325
PR 19-JUL-1999: 99US-0144331
PR 19-JUL-1999: 99US-0144332
PR 19-JUL-1999: 99US-0144333
PR 19-JUL-1999: 99US-0144334
PR 20-JUL-1999: 99US-0144335
PR 20-JUL-1999: 99US-0144352
PR 20-JUL-1999: 99US-0144632
PR 20-JUL-1999: 99US-0144884
PR 21-JUL-1999: 99US-0144814
PR 21-JUL-1999: 99US-0145086
PR 21-JUL-1999: 99US-0145088
PR 22-JUL-1999: 99US-0145089
PR 22-JUL-1999: 99US-0145087
PR 22-JUL-1999: 99US-0145089
PR 22-JUL-1999: 99US-0145192
PR 23-JUL-1999: 99US-0145145
PR 23-JUL-1999: 99US-0145218
PR 23-JUL-1999: 99US-0145224
PR 26-JUL-1999: 99US-0145276
PR 27-JUL-1999: 99US-0145913
PR 27-JUL-1999: 99US-0145918
PR 27-JUL-1999: 99US-0145919
PR 28-JUL-1999: 99US-0145951

PR 02-AUG-1999: 99US-0145986
PR 02-AUG-1999: 99US-0146388
PR 02-AUG-1999: 99US-0146389
PR 03-AUG-1999: 99US-0147038
PR 04-AUG-1999: 99US-0147204
PR 04-AUG-1999: 99US-0147302
PR 05-AUG-1999: 99US-0147192
PR 05-AUG-1999: 99US-0147260
PR 06-AUG-1999: 99US-0147303
PR 06-AUG-1999: 99US-0147416
PR 09-AUG-1999: 99US-0147493
PR 09-AUG-1999: 99US-0147935
PR 10-AUG-1999: 99US-0148171
PR 11-AUG-1999: 99US-0148319
PR 12-AUG-1999: 99US-0148341
PR 13-AUG-1999: 99US-0148565
PR 13-AUG-1999: 99US-0148684
PR 15-AUG-1999: 99US-0149368
PR 17-AUG-1999: 99US-0149175
PR 18-AUG-1999: 99US-0149426
PR 20-AUG-1999: 99US-0149722
PR 20-AUG-1999: 99US-0149723
PR 20-AUG-1999: 99US-0149929
PR 23-AUG-1999: 99US-0149902
PR 23-AUG-1999: 99US-0149930
PR 25-AUG-1999: 99US-0150566
PR 26-AUG-1999: 99US-0150884
PR 27-AUG-1999: 99US-0151065
PR 27-AUG-1999: 99US-0151066
PR 27-AUG-1999: 99US-0151080
PR 30-AUG-1999: 99US-0151303
PR 31-AUG-1999: 99US-0151438
PR 01-SEP-1999: 99US-0151930
PR 07-SEP-1999: 99US-0152363
PR 10-SEP-1999: 99US-0153070
PR 13-SEP-1999: 99US-0153758
PR 15-SEP-1999: 99US-0154018
PR 16-SEP-1999: 99US-0154039
PR 20-SEP-1999: 99US-0154779
PR 22-SEP-1999: 99US-0155139
PR 23-SEP-1999: 99US-0155486
PR 24-SEP-1999: 99US-0155559
PR 24-SEP-1999: 99US-0156458
PR 24-SEP-1999: 99US-0156596
PR 04-OCT-1999: 99US-0157117
PR 05-OCT-1999: 99US-0157753
PR 06-OCT-1999: 99US-0157865
PR 07-OCT-1999: 99US-0158029
PR 08-OCT-1999: 99US-0158232
PR 12-OCT-1999: 99US-0158369
PR 13-OCT-1999: 99US-0159293
PR 13-OCT-1999: 99US-0159294
PR 13-OCT-1999: 99US-0159295
PR 14-OCT-1999: 99US-0159329
PR 14-OCT-1999: 99US-0159330
PR 14-OCT-1999: 99US-0159331
PR 14-OCT-1999: 99US-0159637
PR 14-OCT-1999: 99US-0159638
PR 18-OCT-1999: 99US-0159584
PR 21-OCT-1999: 99US-0160741
PR 21-OCT-1999: 99US-0160767
PR 21-OCT-1999: 99US-0160770
PR 21-OCT-1999: 99US-0160814
PR 21-OCT-1999: 99US-0160815
PR 22-OCT-1999: 99US-0160980
PR 22-OCT-1999: 99US-0160981
PR 22-OCT-1999: 99US-0160989
PR 25-OCT-1999: 99US-0161404
PR 25-OCT-1999: 99US-0161405
PR 25-OCT-1999: 99US-0161406
PR 26-OCT-1999: 99US-0161359
PR 26-OCT-1999: 99US-0161360

PR 26-OCT-1999; 9905-0161361.
 PR 28-OCT-1999; 9905-0161920.
 PR 28-OCT-1999; 9905-0161932.
 PR 28-OCT-1999; 9905-0161993.
 PR 29-OCT-1999; 9905-0162142.

Query Match 81.0%; Score 2951, DB 21, Length 601,
 Best Local Similarity 83.0%; Pred. No. 8,6e-228;
 Matches 567, Conservative 49, Mismatches 75, Gaps 1.

UY 29 LIGSHLDDEKKNVAFKKVKKVKKLDELITVSQVAGIAAANISDVKVELSAKACAKKAS 88
 DB 1 MGSGLDEKKVVEEYRRPVNLSSEELTIGVVAISTVSS--VKVELSTSAVKKAS 58
 UY 89 SIWVESNMKKQTSYGVITGFSATSHRFTKQJAKCKELLPINAGLFGNDETSHTLPH 148
 DB 59 SIWVESNMKKQTSYGVITGFSATSHRFTKQJAKCKELLPINAGLFGNDETSHTLPH 118
 UY 149 SATRAAMIVRINILLQYSGIRFEILEAITKFLNNNITPCLPLKGTITANSQIVPLSYIA 208
 DB 119 SATRAAMIVRINILLQYSGIRFEILEAITSLNNHNSPSPIKGTITANSQIVPLSYIA 178
 UY 209 GLLTRPNSKATGCEVLENAKFAAAGVCGCPPIOPKPCIAIVNCTAVGSGMASMV. 268
 DB 179 GLLTRPNSKATGCEVLENAKFAAAGVCGCPPIOPKPCIAIVNCTAVGSGMASMV. 238
 UY 269 PDANVLLALTEVLSAIFAEVMOCKIPEPTDHLPHKIKHHIGQIFAAALIMFYLLNDSDVKA 328
 DB 239 FPAVQVAVIAEVLASAFAEVMOCKIPEPTDHLPHKIKHHIGQIFAAALIMFYLLNDSDVKA 298
 UY 329 AOKVHEMDEPLAKKQKURKALKTSQWLPPIEVLFSSTKMLBEFINGVNPPILIDVSRNK 368
 DB 299 AOKVHEMDEPLAKKQKURKALKTSQWLPPIEVLFSSTKMLBEFINGVNPPILIDVSRNK 358
 UY 389 ALHGSNPGTPIPIGVSMUDNKLALAIQIKLMAPOHSELVNDYNNGLPSNLSGCHNSLQY 448
 DB 359 ALHGSNPGTPIPIGVSMUDNKLALAIQIKLMAPOHSELVNDYNNGLPSNLSGCHNSLQY 418
 UY 449 GFKGSEIAMAASYCSLQFLANPVINHVSARHONQVNSLQILSAKAVARAVPLIKIMSS 508
 DB 419 GFKGSEIAMAASYCSLQFLANPVINHVSARHONQVNSLQILSAKAVARAVPLIKIMSS 478
 UY 509 TYLVALGOSIDLRHEENKSTVNNTSVQAKKYLINGVNGELHPSPECKDLIRVDPRE 568
 DB 479 TYLVALGOSIDLRHEENKSTVNNTSVQAKKYLINGVNGELHPSPECKDLIRVDPRE 538
 UY 569 VYFAVIDVCSQYTPPLMKILKQVILVDHALNNGELPKNTNTSIFOKIATPEEELKVLIPKE 628
 DB 539 VYFAVIDVCSQYTPPLMKILKQVILVDHALNNGELPKNTNTSIFOKIATPEEELKVLIPKE 598
 UY 629 VHCGRATVENDTLSPNLIKACRSYPLRYFRKELQKGLIQKIVISPGDEFLDKVIFAMC 688
 DB 599 VHAARAAGNGTAPIPNLIKACRSYPLRYFRKELQKGLIQKIVISPGDEFLDKVIFAMC 658
 UY 689 KSGIIDLPLELCSGNSSEPLPIC 711
 DB 659 KSGIIDLPLELCSGNSSEPLPIC 681

RESULT 10

AGS3399 standard; Protein: 736 AA.

AGS3399;

AC AAG53399;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID No: n7482.

XX Protein identification; signal transduction pathway; metabolic pathway;

XX hydridisation assay; genetic mapping; gene expression control; promoter;

XX termination sequence.

OS Arabidopsis thaliana.
 XX EF1033405-A2.
 EN
 XX
 PD 06-SEP-2000.
 XX
 FE 25 FEB-2000; 2000EP-0301439.
 FE 25-FEB-1999; 9905-0139452.
 PR 05-MAR-1999; 9905-0123548.
 PR 09-MAR-1999; 9905-0125788.
 PR 23-MAR-1999; 9905-0126264.
 PR 25-MAR-1999; 9905-0126785.
 PR 29-MAR-1999; 9905-0127462.
 PR 01-APR-1999; 9905-0128234.
 PR 06-APR-1999; 9905-0128714.
 PR 08-APR-1999; 9905-0129845.
 PR 16-APR-1999; 9905-0130077.
 PR 19-APR-1999; 9905-0130449.
 PR 21-APR-1999; 9905-0130510.
 PR 23-APR-1999; 9905-0130891.
 PR 28-APR-1999; 9905-0131449.
 PR 30-APR-1999; 9905-0132048.
 PR 30-APR-1999; 9905-0132407.
 PR 04-MAY-1999; 9905-0132484.
 PR 05-MAY-1999; 9905-0132485.
 PR 06-MAY-1999; 9905-0132486.
 PR 06-MAY-1999; 9905-0132487.
 PR 07-MAY-1999; 9905-0132863.
 PR 11-MAY-1999; 9905-0134256.
 PR 14-MAY-1999; 9905-0134218.
 PR 14-MAY-1999; 9905-0134219.
 PR 14-MAY-1999; 9905-0134221.
 PR 14-MAY-1999; 9905-0134370.
 PR 18-MAY-1999; 9905-0134768.
 PR 19-MAY-1999; 9905-0134941.
 PR 20-MAY-1999; 9905-0135124.
 PR 21-MAY-1999; 9905-0135353.
 PR 24-MAY-1999; 9905-0135629.
 PR 25-MAY-1999; 9905-0136021.
 PR 27-MAY-1999; 9905-0136392.
 PR 28-MAY-1999; 9905-0136782.
 PR 01-JUN-1999; 9905-0137222.
 PR 03-JUN-1999; 9905-0137528.
 PR 04-JUN-1999; 9905-0137502.
 PR 07-JUN-1999; 9905-0137724.
 PR 08-JUN-1999; 9905-0138094.
 PR 10-JUN-1999; 9905-0138450.
 PR 10-JUN-1999; 9905-0138847.
 PR 14-JUN-1999; 9905-0139119.
 PR 16-JUN-1999; 9905-0139452.
 PR 16-JUN-1999; 9905-0139453.
 PR 17-JUN-1999; 9905-0139454.
 PR 18-JUN-1999; 9905-0139457.
 PR 18-JUN-1999; 9905-0139458.
 PR 18-JUN-1999; 9905-0139459.
 PR 18-JUN-1999; 9905-0139460.
 PR 18-JUN-1999; 9905-0139461.
 PR 18-JUN-1999; 9905-0139462.
 PR 18-JUN-1999; 9905-0139463.
 PR 18-JUN-1999; 9905-0139750.
 PR 21-JUN-1999; 9905-0139763.
 PR 21-JUN-1999; 9905-0139817.
 PR 22-JUN-1999; 9905-0139809.
 PR 23-JUN-1999; 9905-0140353.
 PR 23-JUN-1999; 9905-0140354.
 PR 24-JUN-1999; 9905-0140695.
 PR 28-JUN-1999; 9905-0140823.
 PR 29-JUN-1999; 9905-0140991.

Db 360 AATKMEIEELNSVNIINPLIVSEKKAALH33NPWTTPIGVAMINSHIALASIOKIMPAQS 419
QY 424 ELVDEFTNGKPSNLSGGRNPSLUGYGFKGSEIMASTGSEFLPLAFVTHVSAQJHQ 483
Db 420 ELVDEFTNGKPSNLSGGRNPSLUGYGFKGSEIMASTGSEFLPLAFVTHVSAQJHQ 479
QY 484 DVNSIGLSAKRTAEADYDIKIMSSYLVALCOSIDIRHEEMKSTVKNVSOVAKKYL 543
Db 480 DVNSIGLSAKRTAEADYDIKIMSSYLVALCOSIDIRHEEMKSTVKNVSOVAKKYL 539
QY 544 TMGNGELHPSRFCEKDLKRVYDREYFAIIDVCSGTFPLMKLQVLDNALNGETE 603
Db 540 TVGANGELHPSRFCEKDLKRVYDREYFAIIDVCSGTFPLMKLQVLDNALNGETE 599
QY 604 KNTNISIFOKIATFEELKVLIPREVEGRFAVENDTLSTPNRIFKPSYPLPFRVDEL 663
Db 600 ANSATSVHKTGAFEAELKLLPFEVEPRVFEFEITSAIANIKKEPSYPLPFRVDEL 659
QY 664 GRCPITGCKVTSPEERDRVFTAMCKQIIDPLIECLGNGEHP 707
Db 660 NTELTGENVRSPGEEDKVELAISDGLIDPLIECLKEWNGAP 703

RESULT 11
AAG53400
ID AAG53400 standard; Protein; 699 AA.
XX AAG53400:
XX
XX 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment Shu 1b No: 67983.
XX
XX Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EPI033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 9905-0121825.
PR 05-MAR-1999; 9905-0123180.
PR 09-MAR-1999; 9905-0123548.
PR 23-MAR-1999; 9905-0125788.
PR 25-MAR-1999; 9905-0126264.
PR 29-MAR-1999; 9905-0126785.
PR 01-APR-1999; 9905-0127462.
PR 06-APR-1999; 9905-0128234.
PR 08-APR-1999; 9905-0128714.
PR 16-APR-1999; 9905-0129845.
PR 19-APR-1999; 9905-0130077.
PR 21-APR-1999; 9905-0130449.
PR 23-APR-1999; 9905-0130510.
PR 23-APR-1999; 9905-0130891.
PR 28-APR-1999; 9905-0131449.
PR 30-APR-1999; 9905-0132048.
PR 30-APR-1999; 9905-0132407.
PR 04-MAY-1999; 9905-0132484.
PR 05-MAY-1999; 9905-0132485.
PR 06-MAY-1999; 9905-0132486.
PR 06-MAY-1999; 9905-0132487.
PR 07-MAY-1999; 9905-0132863.
PR 11-MAY-1999; 9905-0134286.
PR 14-MAY-1999; 9905-0134218.
PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134221.
PR 14-MAY-1999; 9905-0134370.
PR 18-MAY-1999; 9905-0134768.

PR 19-MAY-1999; 9905-0134941.
PR 20-MAY-1999; 9905-0135124.
PR 24-MAY-1999; 9905-0135353.
PR 25-MAY-1999; 9905-0135629.
PR 27-MAY-1999; 9905-0136021.
PR 28-MAY-1999; 9905-0136392.
PR 01-JUN-1999; 9905-0136782.
PR 03-JUN-1999; 9905-0137222.
PR 04-JUN-1999; 9905-0137528.
PR 07-JUN-1999; 9905-0137502.
PR 08-JUN-1999; 9905-0137724.
PR 10-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 14-JUN-1999; 9905-0138847.
PR 16-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139492.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139763.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139899.
PR 23-JUN-1999; 9905-0140033.
PR 23-JUN-1999; 9905-0140034.
PR 24-JUN-1999; 9905-0140095.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140891.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141842.
PR 02-JUL-1999; 9905-0142154.
PR 06-JUL-1999; 9905-0142055.
PR 08-JUL-1999; 9905-0142390.
PR 08-JUL-1999; 9905-0142803.
PR 09-JUL-1999; 9905-0142820.
PR 12-JUL-1999; 9905-0142977.
PR 13-JUL-1999; 9905-0143542.
PR 14-JUL-1999; 9905-0143624.
PR 15-JUL-1999; 9905-0144005.
PR 16-JUL-1999; 9905-0144085.
PR 16-JUL-1999; 9905-0144086.
PR 19-JUL-1999; 9905-0144325.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144632.
PR 20-JUL-1999; 9905-0144684.
PR 21-JUL-1999; 9905-0144814.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145088.
PR 22-JUL-1999; 9905-0145085.
PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145192.
PR 23-JUL-1999; 9905-0145145.
PR 23-JUL-1999; 9905-0145218.
PR 23-JUL-1999; 9905-0145324.
PR 26-JUL-1999; 9905-0145276.
PR 27-JUL-1999; 9905-0145313.
PR 27-JUL-1999; 9905-0145318.
PR 27-JUL-1999; 9905-0145319.


```

PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157754.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159429.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161461.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0162142.

Query Match 72.78; Score 2648; DR 21; Length 648.
Best Local Similarity 83.4%; Pred No 1 60-203;
Matches 514; Conserved 43; Mismatches 57; Indels 2; Gaps 2.

QY 93 MESMNGCTDGYCTTGCATSHRRKQCGALQKELRPLNACIPRNGT-ETSHRLPHSAT 151
DB 1 MESMNRGTDSYGTGTGATSHRRKQCGALQKELRPLNACIPRNGT-ETSHRLPHSAT 60

QY 152 RAAMIVRINLLDYGSGTGFELIETITKFLNNNTPTCLPRTGTASQPLVLTATGL 211
DB 61 RAAMIVRINLLDYGSGTGFELIETITKFLNNNTPTCLPRTGTASQPLVLTATGL 120

QY 212 TTPPNKAVAPTLEVLNAKATAAAVSGTFFELQFKELALVN-NAVSSMAVLEFA 271
DB 121 TGRPNKAVAPSGTELTJASEAFKLA-V-SGFELQPKELALVNSIAVSSIAVLEFA 179

QY 272 NVLALSLSEVLSAIFARVMGCKPFTDHLTHKIKHHGCTLEAATIMYILDDSDYYKAOK 331
DB 180 NVLALSLSEVLSAIFARVMGCKPFTDHLTHKIKHHGCTLEAATIMYILDDSDYYKAOK 239

QY 332 VHEMDPLQKRPDRAVLRTPQWLSPQTEVFRSTKMFRTFNSVNDPLIVSPNKAHL 391
DB 240 IHEMDPLQKRPDRAVLRTPQWLSPQTEVFRSTKMFRTFNSVNDPLIVSPNKAHL 299

QY 392 GGNUGCTPTVSNMDNTRLAATIGKIMPAQFSELVNDFTNNKLPNSLSSGNRPSTLDYQFK 451
DB 300 GGNUGCTPTVSNMDNTRLAATIGKIMPAQFSELVNDFTNNKLPNSLSSGNRPSTLDYQFK 359

QY 452 GGIAMASYSCTIPLANVINIVHVSAPQHNDVNSIGLISAKTIAFAVDLLKIMSSYVL 511

```

```

DB 360 GAIAMASYSCTIPLANVINIVHVSAPQHNDVNSIGLISAKTIAFAVDLLKIMSSYVL 419
QY 512 VALGUSTDLKHLLEENKSTVKNATVSQVAKVLLIMVNEELHPSKFGEDLLRVDPREVF 571
DB 420 VALGQAVDLHLEENLKKAQKSAVSQVAKRVLTVCANGCHHSRPTEDVLOVDPREVF 479
QY 572 AAYDVGSCYPIPMOKLRQVLVDHALNNGEKKNNISLPFOKIAFEEELKVLIPKEVFG 641
DB 480 SYADDPRLTYPLMOKLRHTLVHDLALDPFANSATSVFIRIGAFEEELKVLIPKEVFG 539
QY 632 VRIAYENDTLSTIPNRKACRSYPLRYFVEELQCHLTQEKVTSQPEFDPVFTAMKGO 691
DB 540 VRIAYEFCISAIANRKYQSSYPLRYFVEELQCHLTQEKVTSQPEFDPVFTAMKGO 599
QY 692 IIDPLEELGAGNGEP 707
DB 600 IIDPLEELGAGNGEP 615

RESULT 13
AAG21947
ID AAG21947 standard; Protein: 423 AA.
AC
AC AAG21947;
XX
DI 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 24692.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
PN
PN HP1033405-A2.
PD
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-030144
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 04-MAP-1999; 99US-0123548.
PR 23-MAP-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134270.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.

```

PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142940.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148665.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150565.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151348.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159283.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160914.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0161484.
PR 25-OCT-1999; 99US-0161484.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

PR 28-OCT-1999: 9905-0161992.
PR 28-OCT-1999: 9905-0161993.
PR 29-OCT-1999: 9905-0162142.

Query Match 51.98; Score 1892; DB 21; Length 423;
Best Local Similarity 83.78; Pred. No. 3,9e-143;
Matches 354; Conservative 33; Mismatches 36; Indels 0; Gaps 0;

```

OY 289 MOSKPEFHDTLHKLKHNGQILAAALMRYLLDGSNYKAAQKRVHMDLQKIKODRAL 348
DB 1 MOSKPEFHDTLHKLKHNGQILAAALMRYLLDGSNYKAAQKRVHMDLQKIKODRAL 60
OY 349 KTSPLWMLSFQIEYRSTKMEFEINSVNDNPLIVSRNKALHSGNFQSTPLGVSMDNTP 408
DB 61 KTSPLWMLSGQIEYRSTKMEFEINSVNDNPLIVSRNKALHSGNFQSTPLGVSMDNTP 120
OY 409 LAIAIAIGKLMFAQFSELVNDFYNNISLPSNMLTASSNPSLDYGFGEKAEIIMASYCELOYL 468
DB 121 LAIAIAIGKLMFAQFSELVNDFYNNISLPSNMLTASSNPSLDYGFGEKAEIIMASYCELOYL 180
OY 469 NPVTNVOSAFQHNQDVNSIGLISAKRTIAEAVDILKMSSTYLVAIGOSIDIRHLEEMK 528
DB 181 NPVTNVOSAFQHNQDVNSIGLISAKRTIAEAVDILKMSSTYLVAIGOSIDIRHLEEMK 240
OY 529 STYKNTVSQVAKKVLTMVNSLELHPSPREKULPEVVIPEVFAVLEOVNSTIPLMKL 588
DB 241 STYKNTVSQVAKKVLTMVNSLELHPSPREKULPEVVIPEVFAVLEOVNSTIPLMKL 300
OY 589 KOVLVVDHALNNGHTEKNTNTSIFOKIATFEELKVLILKEVGVKIAEHNDLSIPNPK 648
DB 301 KOVLVVDHALNNGHTEKNTNTSIFOKIATFEELKVLILKEVGVKIAEHNDLSIPNPK 360
OY 649 ACSQVPLVRFVRELGQFHTGKAKTSDEFEEDRYFTAMCGQILIDPLECIGGNCNPL 708
DB 361 ACSQVPLVRFVRELGQFHTGKAKTSDEFEEDRYFTAMCGQILIDPLECIGGNCNPL 420
OY 709 PTC 711
DB 421 PTC 423

```

RESULT 14
AAB16402
ID AAB16402 standard; Protein: 424 AA.
XX
AC AAB16402;
XX
DT 24-OCT-2000 (first entry)
XX
DE Eucalyptus grandis PAL protein sequence SEQ ID NO. 428
XX
KW plant; lignin biosynthetic pathway; Eucalyptus grandis;
XX
OS pinus radiata; Monterey pine.
XX
XX Eucalyptus grandis.
XX
XX WO200022099-A1.
XX
XX 20-APR-2000.
XX
XX 06-OCT-1999; 9905-NZ00168.
XX
XX 09-OCT-1998; 9905-0164789
XX
XX 14-JUL-1999; 9905-0143811.
XX
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX
XX (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
XX Hloksberg LM, Huvakkala JU.
XX
XX WP1; 2000-317962/27.
XX
XX Novel polynucleotide encoding enzymes involved in lignin-biosynthetic

PT pathway useful for producing transgenic plants especially eucalyptus
PT and pine species having altered lignin content, composition and
PT structure

PS Claim 18, Page 170-171, 213pp, English.

XX The present invention describes isolated polynucleotides and proteins
XX encoding and representing the enzymes cinnamate 4-hydroxylase (C4H),
XX cinnamate 3-hydroxylase (C3H), phenolase (PHL), 6-methyl transferase
XX (OMT), cinnamyl alcohol dehydrogenase (CAD), cinnamoyl-CoA reductase
XX (CCR), phenylalanine ammonia-lyase (PAL), 4-coumarate:CoA ligase (4CL),
XX coniferin glucosyl transferase (CGT), coniferin beta-glucosidase (CBG),
XX laccase, peroxidase, ferulate-5-hydroxylase (F5H), alpha-amylase,
XX caffeic acid methyl transferase, cateoyl CoA methyl transferase,
XX cinnamate CoA ligase, cytochrome P450 1XX1A, diphenol oxidase, flavanol
XX glucosyl transferase, flavonoid hydroxylase, and isoflavone reductase,
XX which are involved in the lignin biosynthetic pathway. The
XX polynucleotides can be used for modulating lignin content, lignin
XX composition and the structure of a plant, especially eucalyptus and pine
XX species, and for modifying the activity of an enzyme involved in lignin
XX biosynthetic pathway, and for producing a plant having altered lignin
XX content, composition and structure. They can be used for designing probes
XX and primers useful for detecting similar DNA and RNA sequences in any
XX organism and for PCR amplification. The lignin content can be efficiently
XX modified using the polynucleotides. AAB7908 to AAB68201 and AAB16341 to
XX AAB16449 represent the polynucleotide and protein sequences used in the
XX exemplification of the present invention.

Sequence 424 AA:

Query Match 49.4%; Score 1799; DB 21; Length 424;
Best Local Similarity 82.68; Pred. No. 1.1e-135;
Matches 351; Conservative 37; Mismatches 29; Indels 8; Gaps 2;

```

OY 1 MENOHNVNGVNNELCT-----KDIPLMNGVAAALATGSHLDYKKKKVAPRRPVKIGCT 55
DB 3 MESTTGTGMLSLISLVAA3SHADPLMVAIAAALITGSHLDEYKPVVEYGPFAVVL3GES 62
OY 56 LTVS0VAGIAAANDSVTVKVELSEARACGVKASSIWMVMSMKGJUSGVJGCGATSHR 115
DB 63 LITIAUYAAYAS--QHGCVGVELSEARAKRPVKASSIWMVMSMKGJUSGVJGCGATSHR 119
OY 116 PTKGSAALCKELIIPUNASIPNNTEITSHLEPSAIPAAIVYINTLLGYSGITFEILE 175
DB 120 RTKGCGALQKELIRPINACIFNGTSCHTLIPGSSIPRAAILVRVNTLLGYSGITFEILE 179
OY 176 AITKPLNNNTTQPLIRGTTASGDIYPLSYLAGLITGPNKSKAGCPGCVLNAKAA 235
DB 180 AITKPLNNHTTCLPLRGTTTASGDIYPLSYTAGLITGPNKSKAVGPGKSLAVEAFRL 239
OY 236 AGVGGFPEIQPEGLALVNGTAVSGMASVILPANNVIALISEVLSAIPAEVWGKPEF 295
DB 240 AGIDTGFPEIQPEGLALVNGTAVSGLASIVLPANNVIALISEVLSAIPAEVWGKPEF 299
OY 296 TDHLTHKLIHHPQIDEAAIMEYLLDGSIDYVKAQKVIHMDPELOKPKODRYALRTSPWL 355
DB 300 TDHLTHKLIHHPQIDEAAIMEYLLDGSIDYVKAQKVIHMDPELOKPKODRYALRTSPWL 359
OY 356 GQDIEVIRKSTIKMEFEINSVNDNPLIVSRNKALHSGNFQSTPLGVSMDNPLIAAIG 415
DB 360 GQDIEVIRKSTIKMEFEINSVNDNPLIVSRNKALHSGNFQSTPLGVSMDNPLIAAIG 419
OY 415 KLMPA 420
DB 420 KLMPA 424

```

RESULT 15
AAB78982
ID AAB78982 standard; Protein: 495 AA.
XX
AC AAB78982;
XX

30-JUL-2002 (first entry)

Lolium perenne LpPALA amino acid sequence SEQ ID No:141.

Lolium perenne; perennial ryegrass; plant; cell wall; lignification; cellulase; enzyme, lignin biosynthesis; cellulose degradation; CCOAMT; caffeoyl-CoA 3-O-methyltransferase; cinammyl alcohol dehydrogenase; CAD; caffeic acid O-methyltransferase; OMT; cinamate-4-hydroxylase; CAH; cinamoyl-CoA reductase; CCR, peroxidase, PER, ferulate-5-hydroxylase; F5H; CELL; phenylalanine ammonia lyase; PAL; 4-coumarate:CoA ligase; 4CL; ryegrass, fescue species; molecular genetic marker.

Lolium perenne.

MO200226994-AL.

04-APR-2002.

28-SEP-2001; 2001WO-AU01221.

29-SEP-2000; 2000AU-0000419.

(AGRI-) AGRIC VICTORIA SERVICES PTY LTD
(AGRE-) AGRSEARCH LTD.

Spangenberg G, Sawbridge TJ, Ong EK, Emmertling M;
WPI: 2002-444025/47.
N-PSDB: ABN87223.

Novel nucleic acid encoding lignification and cellulase enzymes or their related enzymes useful for modifying lignin biosynthesis and cellulose degradation in plants to manipulate plant cell wall

Claim 30: Fig 63; 436pp; English.

The present invention describes a nucleic acid (1) or its fragment encoding caffeoyl-CoA 3-O-methyltransferase (CCOAMT), cinamyl alcohol dehydrogenase (CAD), caffeic acid O-methyltransferase (OMT), cinamate-4-hydroxylase (CAH), cinamoyl-CoA reductase (CCR), peroxidase (PER), cellulase (CELL), ferulate-5-hydroxylase (F5H), phenylalanine ammonia lyase (PAL) or 4-coumarate:CoA ligase (4CL) from perennial ryegrass (Lolium perenne) or fescue species. (1), its nucleotide sequence information and/or single nucleotide polymorphisms is useful as a molecular genetic marker. (1) can be used for modifying lignin biosynthesis and/or cellulose degradation in a plant to manipulate cell walls. (1) or its fragments are useful for isolating cDNAs and genes encoding homologous proteins from the same or other plant species, as hybridisation probes to screen libraries from the desired plant. Short segments of (1) or its fragment are useful in amplification protocols to amplify longer nucleic acids or its fragments encoding homologous genes from DNA or RNA. (1) or its fragments are useful as molecular genetic markers for quantitative trait loci (QTL) tagging, QTL mapping, DNA fingerprinting, and in marker assisted selection, particularly in ryegrass and fescues, and in forage and turf grass improvement, e.g. mechanical stress tolerance, disease resistance, insect pest resistance, plant stature, leaf and stem colour. The present sequence represents a Lolium perenne (perennial ryegrass) amino acid sequence from the present invention.

Sequence 495 AA;

Query Match	48.8%;	Score 1779;	DB 23;	Length 495;
Best Local Similarity	76.6%;	Pred. No. 5.7e-134;		
Matches 344;	Conservative 39;	Mismatches 64;	Indels 2;	Gaps 2

```

OY      263  MAMVLEPDANVIALLESEVISAIFAMVMOGKREFPDHITIKIKHNPQIEAAMIMYIDDG 332
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db       5  LAMVLEPDANVIALLESEVISAIFAMVMOGKREFPDHITIKIKHNPQIEAAMIMYIDDG 64
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      323  SDVVAAAGQVHEMDLQIKQIDVYALKISQWMLSPQIEVIRSSTKMIEKINSVNDPDL 382
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Db	65	XXXXXXIAKKIGCELDPLMKPKQIURXALRISPMQIOPQIUYI RAATKSIRHRI NSVNDPRLI	12
Qy	383	DVSEKNAIHLGNGFUGTPIGVSKMLNTKALAIAGLKFAGFSELYNDFNNKISPLNSGGR	442
Db	125	DVSKGKAIRHGGNFQSTPIGVSKMDTRIALAIGKLFAGFSELYNDFNNKISPLNSGGR	184
Qy	443	NPSLIDGPGGCFIAMAASVCSFLOFLADPVINIHVOSAEOHNOOVNSIGLISARKTAFAVDI	502
Db	185	NPSLIDGGLKGAELAAVASTSELOFGNVTNIHVOSADQIHUVNSIGLISARKTAFAADI	244
Qy	503	LKLMSSYLVALVYGSIDLPHELENNKSTVNTVSAQAKVLTIMVYNGSELHSPFEKEDLL	562
Db	245	LKLMSSYPLVALCOAIDLRHLEFNNRGVAVNCVMAAKRTLTSTNNNGHLMARPCENDLL	304
Qy	563	RVDREYVATIIDVCSGTYTLMQKTLQVLDIADLNNGETENNTSTFQKTAITEEELK	622
Db	305	LTIIDXXAAVAFAVADP-CSTNYPLMGKMAVVEIHALANGCAERDVOTVFARKLAAFEDELK	362
Qy	623	VLLDPIKEVGVRIATENQTLSTINRKACRSTPIYATFVREHLRQGLTGKXVTSIQHTEFLK	682
Db	363	AVLPREVEASAKAAVENGTPDAQONKRTEDCRSTPIYATFVREKXQETYLETDEKTSRQJEEVUK	422
Qy	683	VFTAMCKGQIIDPLLELCEGNNCEPLPIC	711
Db	423	VFTAMKKGXHIIDALLPCLKEKNGPELPIC	451

Search completed. March 29, 2003, 02:10:05
Job time : 71 secs

GenCore version 5.1.4.F5.4578
Copyright (c) 1994 - 2003 CompuGen Ltd

OM protein - protein search, using sw model

Run on: March 29, 2003, 02:09:01, Search time: 39 seconds

(without alignments) 536,403 Million cell updates/sec

Title: US-09-964-992A-1

Perfect score: 3644
Sequence: 1 MENCNHNVCVNNHIAIKPDL... 1100123456789101112131415161718192021222324252627282930313233343536373839404142434445

Scoring table: BLASTNM2

Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: Issued_Patents_AA*

1: /cgn2_6/prodata/1/1aa/5A.COMB.pep.*
2: /cgn2_6/prodata/1/1aa/5H.COMB.pep.*
3: /cgn2_6/prodata/1/1aa/6A.COMB.pep.*
4: /cgn2_6/prodata/1/1aa/6B.COMB.pep.*
5: /cgn2_6/prodata/1/1aa/6C.COMB.pep.*
6: /cgn2_6/prodata/1/1aa/6D.COMB.pep.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1799	49.4	424	4	US-09-615-192A-328
2	1345	36.9	412	4	US-09-615-192A-330
3	1025	28.1	720	4	US-09-624-693A-13
4	994.5	27.3	716	4	US-09-624-693A-19
5	994.5	27.3	716	4	US-09-627-216A-8
6	991.5	27.2	716	4	US-09-627-216A-10
7	971.5	26.7	713	4	US-09-624-693A-17
8	967	26.5	726	4	US-09-624-693A-21
9	939	25.8	726	4	US-09-615-192A-327
10	844.5	23.2	740	4	US-09-624-693A-15
11	653	17.9	164	4	US-09-615-192A-325
12	444	11.4	114	4	US-09-615-192A-326
13	384	10.7	97	4	US-09-615-192A-329
14	276	7.6	132	4	US-09-615-192A-331
15	122.5	3.4	2411	4	US-09-624-693A-36
16	119.5	3.3	1912	4	US-08-409-995-4
17	119.5	3.3	1912	4	US-08-685-467-4
18	119.5	3.3	2353	4	US-09-377-155-33
19	119.5	3.3	2353	4	US-08-913-942-4
20	119.5	3.3	2353	4	US-09-669-974-33
21	119.5	3.3	2354	4	US-09-268-347-47
22	118.5	3.3	934	1	US-08-215-805A-80
23	118.5	3.3	2385	4	US-09-308-375-2
24	117	3.2	683	4	US-09-620-415C-157
25	116	3.2	831	4	US-09-346-447A-1
26	116	3.2	1776	4	US-09-556-877-179
27	116	3.2	1776	4	US-09-620-412C-179

28	113.5	3.1	884	4	US-09-741-150-4	Sequence 4, App1
29	110.5	3.0	1500	4	US-09-323-472A-2	Sequence 2, App1
30	110.5	3.0	1500	4	US-09-323-472A-12	Sequence 4, App1
31	110.5	3.0	1500	4	US-09-323-472A-14	Sequence 12, App1
32	110.5	3.0	1500	4	US-09-323-472A-14	Sequence 14, App1
33	110	3.0	2048	4	US-09-268-347-48	Sequence 48, App1
34	109	3.0	1022	4	US-08-772-270A-2	Sequence 2, App1
35	109	3.0	10182	4	US-09-134-061C-3159	Sequence 3159, App1
36	108.5	3.0	849	4	US-09-157-257-4	Sequence 4, App1
37	107.5	3.0	1151	4	US-09-134-001C-3242	Sequence 4, App1
38	107.5	3.0	1833	4	US-08-621-944A-4	Sequence 4, App1
39	107.5	3.0	1833	4	US-08-945-567D-4	Sequence 4, App1
40	107.5	3.0	1992	4	US-08-621-944A-3	Sequence 3, App1
41	107.5	3.0	1992	4	US-08-945-567D-3	Sequence 3, App1
42	106.5	2.9	548	2	US-08-467-822-31	Sequence 31, App1
43	106.5	2.9	548	4	US-08-472-647-31	Sequence 31, App1
44	106.5	2.9	548	4	US-08-466-248-31	Sequence 31, App1
45	105.5	2.9	730	4	US-09-398-865A-2	Sequence 2, App1

ALIGNMENTS

RESULT 1
US-09-615-192A-328
Sequence 328, Application US/09615192A
Patent No. 6410718
GENERAL INFORMATION:
APPLICANT: Novokkale, Leonard N.
APPLICANT: Novokkale, Ilkka
TITLE OF INVENTION: Materials and Methods for the
FILE REFERENCE: 11090.1093640
CURRENT APPLICATION NUMBER: US/09/615,192A
CURRENT FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 08/975,316
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: US 08/713,000
PRIOR FILING DATE: 1996-09-11
PRIOR APPLICATION NUMBER: US 09/169,789
NUMBER OF SEQ ID NOS: 405
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 328
LENGTH: 424
TYPE: PRT
ORGANISM: Eucalyptus grandis
US-09-615-192A-328

Query Match 49.4%; Score 1799; DB 4; Length 424;
Best Local Similarity 82.6%; Pred. No. 3e-154;
Matches 351, Conservative 37, Mismatches 29, Indels 8, Gaps 2;

US-09-615-192A-328
116 KIKGQGLAKLKLKPLKINAGIPLGNGTTSHTPLHSATRAAMIVRITLQCGSGIRPELLE 175
116 KIKGQGLAKLKLKPLKINAGIPLGNGTTSHTPLHSATRAAMIVRITLQCGSGIRPELLE 175
120 FTKGKALQKELLPLFNAIIPFNSTFCHTLPLSTPAAMIVRNTLQCGSGIRPELLE 179
120 FTKGKALQKELLPLFNAIIPFNSTFCHTLPLSTPAAMIVRNTLQCGSGIRPELLE 179
176 ATRKPLNNITPLDPLKGTITASGIVPLSYIAGILGCPNSKAVGPGIVINAKAPAA 235
176 ATRKPLNNITPLDPLKGTITASGIVPLSYIAGILGCPNSKAVGPGIVINAKAPAA 235
180 ATRKPLNNITPLDPLKGTITASGIVPLSYIAGILGCPNSKAVGPGIVINAKAPAA 239
180 ATRKPLNNITPLDPLKGTITASGIVPLSYIAGILGCPNSKAVGPGIVINAKAPAA 239
236 ASVEGHEFLQPEKGLALVNGTAVSGMAKSVLFEANVLALESEVLSAIFAEVMGKEFE 295
236 ASVEGHEFLQPEKGLALVNGTAVSGMAKSVLFEANVLALESEVLSAIFAEVMGKEFE 295
240 AGHTGCFHLPQKGLALVNGTAVSGMAKSVLFEANVLALESEVLSAIFAEVMGKEFE 299
240 AGHTGCFHLPQKGLALVNGTAVSGMAKSVLFEANVLALESEVLSAIFAEVMGKEFE 299

RESULT 3
US-09-624-693A-13
; Sequence 13, Application US/09624693A
; Patent No. 6355468
; GENERAL INFORMATION:

RESULT 4
US-09-624-693A-19
; Sequence 19, Application US/09624693A
; Patent No. 6355468

GENERAL INFORMATION:
APPLICANT: Yoshida, Roberta
TITLE OF INVENTION: Phenylalanine Ammonia lyase Polypeptide and
TITLE OF INVENTION: Polypeptide Sequences and Methods of Obtaining and
TITLE OF INVENTION: Using Same
FILE REFERENCE: 29479/500NSC
CURRENT APPLICATION NUMBER: US/09/624,693A
CURRENT FILING DATE: 2000-07-24
NUMBER OF SEQ ID NOS: 25
SOFTWARE: Patent Ver. 2.0
SEQ ID NO: 19
LENGTH: 716
TYPE: PRT
ORGANISM: Rhodotorula toruloides
US-09-624-693A-19

Query Match 27.3% Score 994.5; Dh 4; Length 716;
Best Local Similarity 35.9%; Pred. No. 4,9e-81;
Matches 270; Conservative 123; Mismatches 250; Indels 109; Gaps 23;

QY 8 NGVNNELCTKIDPLNMCVAAEALTGSHL-----DEKKKVAERFRKIVVKGLETTLS 59
DB 14 NGVAS---AKQAVNGASTNLAAGSHLPTQVGYDVIKEMLAAPDSTLIDGYSINIG 70
QY 60 QVAGTAAANDSDTVKVELSEARAGYKASSDVMESNMKCTDSCVTTCGATSHRTKO 119
DB 71 DV--VSAARKGRPVYRKSDDELRSKIDKSVETLSQLSMSV--YVITTCGASADTPTED 126
QY 120 GCAIQELIRPLNAGI-----FCNGTETSHTLPHSATRAAMIVKINTLIDGYSIR 170
DB 127 AISTQALIEHQLGCVLPSFISFRLGRLENS--LPLEVVGAMTIVNLSLTKHSAYR 184
QY 171 FELLEATFELNNITPCLELTITASGDLVPLSYAGLTGRLNSKA--VGPG--EVL 227
DB 185 LVVLEATFELNNITPCLELTITASGDLVPLSYAGLTGRLNSKA--VGPG--EVL 227
QY 228 NAKKAAAGVGEFPELOPKKGLAVNGTAVSGMASMVLFDANVLAISVLSALFAE 287
DB 245 YAREMALNLEP--VLCPRKICGLVNGTAVSASMATLALHDAMLSLSOSTTAMTVE 302
QY 288 VMGKRP-EFTDHLTHKLIKHHPGQIEAALIMFYILDGSDY--VKAOKVHEMDPIQPKODR 345
DB 303 AMVGHGSHFPLDHTVTRHPDQIEVAGNIRKLKLSGSRFAVHHEFFVVKDDEGLRDDR 362
QY 346 YALPTSPQMLPGDI-EVIPSSTKMIEREINSVNDPLIDVSNKALHGCNFGTPTGVSM 404
DB 363 YPLPTSPQMLPGDI-EVIPSSTKMIEREINSVNDPLIDVSNKALHGCNFGTPTGVSM 404
QY 405 DNTRLAIAAIGKIMFQFSELYNDYVNGIJSNLSGGRNPSIDYFGKGEIAMASVCSSEL 464
DB 423 EKTRRLGIAQIGKINFTQLTLEMNAGNNRGLPSCLA-AEDPISLHYCKGLIDIAAAVYSEL 481
QY 465 QPLANVTNHHVOSAPHOHNDVNSICLISAKRTAFAYVDIILKMSSTYVLAICOSTIDRLHE 524
DB 482 GHLANVTNHHVOPAEWANOAVNSIALISAKRTTESNDVLSILLATHLCVLAIDILRLAE 541
QY 525 ENNKSTVKNTVSOVAKKVLITMGVNGELIHSRCEKDLLRVVDREVFAVTDVNSGTYPL 584
DB 542 FEFKK-----QFGPA-----IVSLIDQHF-----GSAMT 565
QY 585 MOKLROVLVDHALNNGETKNTNTSIFOKIATFEELKVLIPKVEGV----- 632
DB 566 GSNLKRDELVE-KVNTKLARLEDTNSYDLPVPMIDAFSEMASTVEVLSLSLAAVNA 624
QY 633 -RIAYENDITSLINRIK-----ACKSYF-----LYRFRREIG-----RG--FL 668
DB 625 WKMAAASALSLIKVYRRETFMSASTSSVLSYLSIKRTUULIYAFVKHIDYAKARKQDYFL 684
QY 669 TGEKVTSPDEEDRVFTAMCKGQIIDPLECL 700
DB 685 GKOHVTL-GSNVSKLYEALIKSGHINNVLKML 715

RESULT 5

US-09-627-216A-8
Sequence 8, Application US/09627216A
Patent No. 6368837

GENERAL INFORMATION:

APPLICANT: Sartisanti, Sima F
APPLICANT: Tang, Xiao-Song
APPLICANT: Qi, Wei-Wei
APPLICANT: Vannelli, Todd
APPLICANT: Galenby, Anthony
TITLE OF INVENTION: Bioproduction of para-Hydroxycinnamic Acid
CURRENT APPLICATION NUMBER: US/09/627,216A
CURRENT FILING DATE: 2000-07-27
PRIOR APPLICATION NUMBER: 60/147,719
PRIOR FILING DATE: 1998-08-06
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Microsoft Office 97
SEQ ID NO: 8
LENGTH: 716
TYPE: PRT
ORGANISM: Rhodotorula glutinis
US-09-627-216A-8

Query Match 27.3% Score 994.5; Dh 4; Length 716;
Best Local Similarity 35.9%; Pred. No. 4,9e-81;
Matches 270; Conservative 123; Mismatches 250; Indels 109; Gaps 23;

QY 8 NGVNNELCTKIDPLNMCVAAEALTGSHL-----DEKKKVAERFRKIVVKGLETTLS 59
DB 14 NGVAS---AKQAVNGASTNLAAGSHLPTQVGYDVIKEMLAAPDSTLIDGYSINIG 70
QY 60 QVAGTAAANDSDTVKVELSEARAGYKASSDVMESNMKCTDSCVTTCGATSHRTKO 119
DB 71 DV--VSAARKGRPVYRKSDDELRSKIDKSVETLSQLSMSV--YVITTCGASADTPTED 126
QY 120 GCAIQELIRPLNAGI-----FCNGTETSHTLPHSATRAAMIVKINTLIDGYSIR 170
DB 127 AISTQALIEHQLGCVLPSFISFRLGRLENS--LPLEVVGAMTIVNLSLTKHSAYR 184
QY 171 FELLEATFELNNITPCLELTITASGDLVPLSYAGLTGRLNSKA--VGPG--EVL 227
DB 185 LVVLEATFELNNITPCLELTITASGDLVPLSYAGLTGRLNSKA--VGPG--EVL 227
QY 228 NAKKAAAGVGEFPELOPKKGLAVNGTAVSGMASMVLFDANVLAISVLSALFAE 287
DB 245 YAREMALNLEP--VLCPRKICGLVNGTAVSASMATLALHDAMLSLSOSTTAMTVE 302
QY 288 VMGKRP-EFTDHLTHKLIKHHPGQIEAALIMFYILDGSDY--VKAOKVHEMDPIQPKODR 345
DB 303 AMVGHGSHFPLDHTVTRHPDQIEVAGNIRKLKLSGSRFAVHHEFFVVKDDEGLRDDR 362
QY 346 YALPTSPQMLPGDI-EVIPSSTKMIEREINSVNDPLIDVSNKALHGCNFGTPTGVSM 404
DB 363 YPLPTSPQMLPGDI-EVIPSSTKMIEREINSVNDPLIDVSNKALHGCNFGTPTGVSM 404
QY 405 DNTRLAIAAIGKIMFQFSELYNDYVNGIJSNLSGGRNPSIDYFGKGEIAMASVCSSEL 464
DB 423 EKTRRLGIAQIGKINFTQLTLEMNAGNNRGLPSCLA-AEDPISLHYCKGLIDIAAAVYSEL 481
QY 465 QPLANVTNHHVOSAPHOHNDVNSICLISAKRTAFAYVDIILKMSSTYVLAICOSTIDRLHE 524
DB 482 GHLANVTNHHVOPAEWANOAVNSIALISAKRTTESNDVLSILLATHLCVLAIDILRLAE 541
QY 525 ENNKSTVKNTVSOVAKKVLITMGVNGELIHSRCEKDLLRVVDREVFAVTDVNSGTYPL 584
DB 542 FEFKK-----QFGPA-----IVSLIDQHF-----GSAMT 565
QY 585 MOKLROVLVDHALNNGETKNTNTSIFOKIATFEELKVLIPKVEGV----- 632
DB 566 GSNLKRDELVE-KVNTKLARLEDTNSYDLPVPMIDAFSEMASTVEVLSLSLAAVNA 624

QY	633	RAYANDLISLPNPK-----ACRSP-----LYRFRLEELG-----RG-FL 668
Db	625	MKVAALSAISLIRQVRETFMSAASSTSPALSTLSPRTQILVAFRRELCAKARGCVFL 684
QY	669	TGCKVITSPGEEDRYFTAMCCQLIDPLECL 700
Db	685	GKQEVIT-GSNVSKITAEIKSGRIINNVLMKL 715

RESULT 6
US-09-627-216A-10

```

; OTHER INFORMATION: Description of Artificial Sequence: mutant from
; OTHER INFORMATION: Rhodotorula glutinis
US-09-627-216A-10

```

[illegible]

```

RESULT 7
US-09-624-693A-17
: Sequence 17, Application US/09624693A
: Patent No. 6355468
: GENERAL INFORMATION:
: APPLICANT: Yoshida, Roberta
: APPLICANT: Koobstra, Anna
: TITLE OF INVENTION: Phenylalanine Ammonia Lyase Polypeptide and
: TITLE OF INVENTION: Polynucleotide Sequences and Methods of Obtaining and
: TITLE OF INVENTION: Using Same
: FILE REFERENCE: 294 79/500NSC
: CURRENT APPLICATION NUMBER: US/09/624,693A
: CURRENT FILING DATE: 2000-07-24
: NUMBER OF SEQ ID NOS: 25
: SOFTWARE: patentln Ver. 2.0
: SEQ ID NO 17
: LENGTH: 713
: TYPE: PRT
: ORGANISM: Rhodotorula mucilaginosa
US-09-624-693A-17

```



```

QY 407 TRLAIAIGKLMFAESELVNFENNNLPSLNSGIRNPSITFYFKGJE:ELAMASYSELDF 466
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 427 TRIAVALMKRQVSTQI:THHLMNAGMNPALPSCIA--AED:SLSTHCKGIDIAAAATSELGH 485
QY 467 LANPTNHOVAOHNOVDNLSGLISAKRTARAVDILKIMSTYVALCOSIDLRLHLEN 526
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 486 LANPSTHVOAPAMNSQVAINSLATLSARTRAEANDVLNLLATILHYVLQVAVDELRAMEFE 545
QY 527 MKSTVNTYVQAKKYLIMGVNGBLHPSHCKKULLKRVDRKYVAYIDVCSGTYLIMU 586
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 546 HTKAFPMVTEILKQHF-----GALATAHVPDK-----VRKSTYKRIQO--NNSYDLEQ 592
QY 587 KLRQVL-----VDILNNGEETKNTNTSIFOKIATPEEELKVLPLKREVEYRIAYENDT 640
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 593 RMDHTSVATGAVVIALAQGVSLASIAN--WKVACAKKALAL-----TRSGYKSTPMAP 645
QY 641 LSLPNRIK--ACKSYLYKPVHREHG---KG--FLTCBKVTSPEGRDHYVTAMCKQOI 692
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 646 SSSSPALKYLSPTFVLTSFVREEVYVKARPGCVLYIKGEVTI GTNVSRIYFAIKSR:1 704
QY 693 IDPLILPCL 700
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 705 APVILVKKM 712

```

RESULT 8

```

US-09-624-693A-21
: Sequence 21, Application US/09624693A
: Patent No. 6353468
: GENERAL INFORMATION:
: APPLICANT: Yoshida, Roberta
: APPLICANT: Koolstra, Anna
: TITLE OF INVENTION: Polynucleotide Ammonia lyase polypeptide and
: TITLE OF INVENTION: Polynucleotide Sequences and Methods of Obtaining and
: TITLE OF INVENTION: Using Same
: FILE REFERENCE: 29479/500NSC
: CURRENT APPLICATION NUMBER: US/09/624-693A
: CURRENT FILING DATE: 2000-07-24
: NUMBER OF SEQ ID NOS: 25
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 21
: LENGTH: 726
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: NAME/KEY: SITE
: LOCATION: 112, 16, 17, 20, 24, 25, 28, 30, 38, 42, 47, 48, 56, 57, 62, 66,
: OTHER INFORMATION: "Xaa" means any amino acid; "Xaa" means no consensus at that
: OTHER INFORMATION: position
: OTHER INFORMATION: Description of Artificial Sequence: Consensus of
: OTHER INFORMATION: SEQ ID NOS: 13, 17, and 19
US-09-624-693A-21

```

Query Match

26.5%: Score 967; DP 4; Length 726;

Best Local Similarity 36.9%: Pred No. 1 5e-78;

Matches 263; Conservative 108; Mismatches 262; Indels 80; Gaps 20;

```

QY 32 SHIDPEYKMWAF-FKKPVKICQETITVSOVAGIAAANDSDIKVYLSAAGACVAKSSD 90
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 49 TOLDIYEXXKADPXTDDXXELGYSLTLDVYG--AAKRGVRRVXKXDDFTKXDKQKVE 106
QY 91 WVMESNKKGTSTYVTTGRTATSHRFTKQJALQKELIFELNAGI-----FGNGTE 141
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 107 FLKXQJXNSV--YGVITGFGSADRTEDAIISLOKALIEHQICVAPISXDSFXLGRGLK 164
QY 142 TSTIDHSATRAMVIRINTLLQVSGIRFELLEITKTLNNTTPCPPLRFITIASGDL 201
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 165 NS--LPLEYVFGAMTIRVNSITRGISAVRLVLEALTNLNGITPIVPLRGTISASGDL 222
QY 202 VPLSYIAGLITGPNNSKA--VGPFG--FVILNAKKAPAAACVPGCPPELQPKKGLAVNGTA 258
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 223 SPLSYIAAALITIFGSKVIVIHGEXEKIMXAHFAIALFLER--VVL:PKKEL:SLVNITA 280

```

```

QY 459 VSGMASMYLFTANVALLSEVLSAIFAEMQKRP-ETTEHLTKRLKHGQIEAAME 317
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 281 VSASMTALIAHAAHMLISLSOAITALTAYAMAGASHPELIMVTRDPQILIAVNRIR 340
QY 318 YLIDSDUY--VKAQKQVHEMDPIQKQKQBYAITSPOWICPOI--EVIHSSKMIEREINS 375
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 341 TLEDSXFPVHHEEVEKVKDDEGILKQDKYPLKTSPOWLSPLVSNMILIAVLELNGUS 400
QY 376 VINDPILIVSRKALHGCNPHOCTPIGVSMONTKILAAIAIKIMRAGSELVNDVYNNQIP 435
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 401 TTNPPLIDVENKXTHHGGNMFQASAVXNTEKTRIALALIGKINLPQI:EMN:NAOMNHOIP 460
QY 436 SNISGGRNSIDYGPGR:ETAMASYSELDFLANVTNIVGSAEJHN:DVNSLISAR 495
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 461 SCIA--AED:SLSTHCKGIDIAAAATSELGLANPVTTHVOIALHNGOAVSLAISAR 519
QY 496 TARAVDILKIMSTYVALCOSIDLRLHLENKSTVKNVSOVAKKYLIMGVNGBLHPS 555
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 520 IAVANVLNLLATILHYVLQVAVDELRAMEFEKKQFPPXXXXLXQIF---GXXXTXX 574
QY 556 FCKKDLKRVDRKYVAYIDVCSGTYPIMOKLROYLVDAHANGETKNTNTSIFOKIA 615
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 575 XXXXELXXKVVXX-----LKKRLQOT-----NSYDLRPMHDAFSXATG 613
QY 616 IFEEELK-----VLLPEVEGVRIAYENDTSLSPRIK-----ACKSY 653
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 614 TVEXELSSXXXXXVSLAAVANAKVAAKKAISLTXVXXKXXVXAAISSSPALXISPRIR 673
QY 654 PLVPEVREHG---KG--FLTCBKVTSPEGRDHYVTAMCKQOI:IDPLILPCL 700
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 674 VLVSFVEELGVAPRGVFLKQEVTI--GSNVSRIYFAIKSGPINXVLVYML 725

```

RESULT 9

```

US-09-615-192A-327
: Sequence 327, Application US/09615192A
: Patent No. 6410718
: GENERAL INFORMATION:
: APPLICANT: Bloksberg, Leonard N.
: APPLICANT: Havukkala, Jukka
: TITLE OF INVENTION: Materials and Methods for the
: TITLE OF INVENTION: Modification of Plant Lipid Content
: FILE REFERENCE: 11000 1003c40
: CURRENT APPLICATION NUMBER: US/09/615-192A
: CURRENT FILING DATE: 2000-07-12
: PRIOR APPLICATION NUMBER: US 08/975,316
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: US 08/714,000
: PRIOR FILING DATE: 1996-04-11
: PRIOR APPLICATION NUMBER: US 09/164,789
: PRIOR FILING DATE: 1998-10-09
: NUMBER OF SEQ ID NOS: 405
: SOFTWARE: Blast2 for Windows Version 3.0
: SEQ ID NO 327
: LENGTH: 226
: TYPE: PRT
: ORGANISM: Eucalyptus grandis
US-09-615-192A-327

```

Query Match

25.8%: Score 939; DP 4; Length 226;

Best Local Similarity 80.5%: Pred No. 6 9e-77;

Matches 182; Conservative 19; Mismatches 25; Indels 0; Gaps 0;

```

QY 444 PSIDYCKGKGLIAMASYSELDFLANPVTNIVHVSARHNOVDNLSGLISAKRTARAVDIL 503
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 1 PSIDYCKGKGLIAMASYSELDFLANPVTNIVHVSARHNOVDNLSGLISAKRTARAVDIL 60
QY 504 KLMSTYVALCOSIDLRLHLENKSTVKNVSOVAKKYLIMGVNGBLHPSRFEKDLK 563
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 61 KLMSTYVALCOSIDLRLHLENKSTVKNVSOVAKKYLIMGVNGBLHPSRFEKDLK 120
QY 564 VLVREVFVAVILGVNSGIVPLMKLROYLVDAHANGETKNTNTSIFOKIATPEEELK 623
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

```

DB 121 VYREYFAYIDUP:SAIYPLMOKLROVLVDALLDVEKNEKSTISFOKIGAEFEELKA 180
 QY 624 ILPKVEGVRIAYENDTLSTIPNRIKACRSYPIYFVREHIGRGFLT 669
 DB 181 LIPKVEYENAKRQFESGNSAIANKIRGCRSYPLIFVREHIGRGFLT 226

RESULT 10
 US-09-624-693A-15
 : Sequence 15, Application US/09624693A
 : Patent No. 6355468
 : GENERAL INFORMATION:
 : APPLICANT: Yoshida, Roberta
 : APPLICANT: Kootstra, Anna
 : TITLE OF INVENTION: Phenylalanine Ammonia Lyase Polypeptide and
 : TITLE OF INVENTION: Polynucleotide Sequences and Methods of Obtaining and
 : TITLE OF INVENTION: Using Same
 : FILE REFERENCE: 29479/500NSC
 : CURRENT APPLICATION NUMBER: US/09/624.693A
 : CURRENT FILING DATE: 2000-07-24
 : NUMBER OF SEQ ID NOS: 25
 : SOFTWARE: Patent In Ver. 2.0
 : SEQ ID NO 15
 : LENGTH: 740
 : TYPE: PRT
 : ORGANISM: Amanita muscaria
 US-09-624-693A-15

Query Match 23.2%, Score 844.5; DB 4; Length 740;
 Best Local Similarity 31.3%; Pred. No. 2e-67;
 Matches 229; Conservative 153; Mismatches 266; Indels 83; Gaps 21;

QY 19 PLNMGVAABALT-----GSHDEVKKVAAPFRKPVNLGSEPTLVSOVAGIAA 66
 DB 27 PVN-GFKATALSKSRMTKTSTALSOFLAYRELEGYKNGRAIKVDQTLIAAVA-AA 83
 QY 67 ANDSTVVELSEARAGVAKASDWMVMKNTDSYVITGREGATSHRETKUGSALQKE 126
 DB 84 ARYNAVEIDESPVLKVEYKKSQALAIANKSTGASVGLSTGFGSGADTRDKMILGFA 143
 QY 127 LIRPILACIFGNGTR-----TSHLLPHSATRAAMIVRINILAQYSGIRFELIEA 176
 DB 144 ILQHGVLSTPTEFLVDLPLODANNTSMPEAMIRGAILIRMSLLIRHSGIRWELIEK 203
 QY 177 ITRPLANNITPLRGITITASGILVPISTYAGILTCRPNKAV-GPT---CHVINAER 231
 DB 204 MEHLIAAVYIVVPIRGSISSGDISPLSYIAGTIIIGNPSIKYVHGPKSGQIRIGSSKD 263
 QY 233 AFNAAGVEGFELOPKFEGIALVNGIANGSGMASVLEFDANVLLALSEVLSATFAEYMOG 291
 DB 264 VIALHNIEP--FPLESKREPLGLINGTASASVAAALALNEALHVLVLAQVCTAMGTALIG 321
 QY 292 KPEFTDLTHLKHHPQSOFAAIMEYILDSIDYKAAQ-KVHEM--DPLQPKADRY 346
 DB 322 TRASHAPFIHATARPHRPGVECAENIMWLDGS---KLAQEEHEVLEDDKYLTRQDRY 378
 QY 347 ALRTSPOMLQOIVYIRSTKMIKEIN--SVNPNLLIDWSKNNALHGCNQGPIPVSM 404
 DB 379 PLRTSPQELQOIFDIISAFQTVTOECVYLPATDNPILIDGTGSHHGQNFQAMAVYNAM 438
 QY 405 ENTKLALAIQKLMFAJSELVNDFYNNGLPESNLGGRNSLSDGFGKGLAMASYCSEL 464
 DB 439 EKTRLALHVGKLLFSQSTELVNPAMNKGILPPSVA-ANDPSLNAHAGLIDIAAAYAE- 496
 QY 465 OFLAPVIVNHVQASQOHQDYNISLISAKKTAAVILLKLMSTYVALICQSIDIRHLE 524
 DB 497 --ATPGTHIISQEMHQAVNSLISARATITSLLEVLSLISASYLIICQALDLIALQ 553
 QY 525 ENMKSTVANTYSQVAKKYLIMGVNGELAPS--RPECMDLLRVYDREYVAYIDYQSGT- 581
 DB 554 RFFLPG-----LDLIIIEELRSSGSLSSQEMKLOONLTSAEHDLKITT 601
 QY 582 -----YPLMOKLROVLVDHALNNGETEKNTSIFOKIATFEELKVLPRVEGVRI 634

DB 602 MONTORMTMATSSSVLLQFDTDSGASVPPSSODLLSSVSSQSSVATRSSVILMDLK 661
 QY 635 AY-ENMTIISINWIKACRSYPIYFVREHIG-----FGLTIG---KRVISQGEHD 681
 DB 662 EYIFGDSPTPTPASQYIKRTPPVYQFIPITISVGRKHSSENYNKKFYNSLVEDVTI-GQNSI 720
 QY 682 RYETAMCKGOI 692
 DB 721 RYIESIRQCKM 731

RESULT 11
 US-09-615-192A-325
 : Sequence 325, Application US/09615192A
 : Patent No. 6410718
 : GENERAL INFORMATION:
 : APPLICANT: Bloksberg, Leonard N.
 : APPLICANT: Havukkala, Ilkka
 : TITLE OF INVENTION: Materials and Methods for the
 : TITLE OF INVENTION: Modification of Plant Lignin Content
 : FILE REFERENCE: 11000.1003c4U
 : CURRENT APPLICATION NUMBER: US/09/615.192A
 : CURRENT FILING DATE: 2000-07-12
 : PRIOR APPLICATION NUMBER: US 08/975.316
 : PRIOR FILING DATE: 1997-11-21
 : PRIOR APPLICATION NUMBER: US 08/713.000
 : PRIOR FILING DATE: 1996-09-11
 : PRIOR APPLICATION NUMBER: US 09/169.789
 : NUMBER OF SEQ ID NOS: 405
 : SOFTWARE: FastSeq for Windows Version 3.0
 : SEQ ID NO 325
 : LENGTH: 164
 : TYPE: PRT
 : ORGANISM: Eucalyptus grandis
 US-09-615-192A-325

Query Match 17.9%, Score 653; DB 4; Length 164;
 Best Local Similarity 73.6%; Pred. No. 3e-51;
 Matches 120; Conservative 16; Mismatches 27; Indels 0; Gaps 0;

QY 548 NEELHPSRCECKDLLRYVDREYFAYIDVCSYPIYFVREHIGRGFLT 607
 DB 1 NEELHPSHYCERDLKVVDPHFVYANDGSAIYPLMOKLROVLVDQALVNGESHLNPS 60
 QY 608 TSIFQKIATFEELKVLPRKVEGVRIAYENDTLSTIPNRIKACRSYPIYFVREHIGRG 667
 DB 61 TSIFQKIATFEELKVLPRKVEGVRIAYENDTLSTIPNRIKACRSYPIYFVREHIGRG 667
 QY 668 LTGKVTSPGEEFDPVFTAMCKGOIIDLPLECJGANGGEPLPT 710
 DB 121 LTGKVTSPGEEFDPVFTAMCKGOIIDLPLECJGANGGEPLPT 163
 RESULT 12
 US-09-615-192A-326
 : Sequence 326, Application US/09615192A
 : Patent No. 6410718
 : GENERAL INFORMATION:
 : APPLICANT: Bloksberg, Leonard N.
 : APPLICANT: Havukkala, Ilkka
 : TITLE OF INVENTION: Materials and Methods for the
 : TITLE OF INVENTION: Modification of Plant Lignin Content
 : FILE REFERENCE: 11000.1003c4U
 : CURRENT APPLICATION NUMBER: US/09/615.192A
 : CURRENT FILING DATE: 2000-07-12
 : PRIOR APPLICATION NUMBER: US 08/975.316
 : PRIOR FILING DATE: 1997-11-21
 : PRIOR APPLICATION NUMBER: US 08/713.000
 : PRIOR FILING DATE: 1996-09-11
 : PRIOR APPLICATION NUMBER: US 09/169.789

NUMBER OF SEQ ID NOS: 405
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 326
LENGTH: 114
TYPE: PR
ORGANISM: Eucalyptus grandis
US-09-615-192A-326

Query Match 11.9%: Score 434; DB 4; Length 114;
Best Local Similarity 71.1%: Pred. No. 1e-11;
Matches 81; Conservative 12; Mismatches 21; Indels 0; Gaps 0;

QY 592 LVHAIINNGEIKNTSTFQKIAIEFEELKVLKFEVGVIAVENDTSLPNRIKACR 651
DB 1 LVDAQVNESEINPSTSTFOKIVAFEEELKVLKFEVGVIAVENDTSLPNRIKACR 60
QY 652 SYPLVYFVEDELGPGFTTGEKVTSPGEFTVFTAM*KGU110PLEL*LVSMNG 705
DB 61 SYPLVYKIVREHLOTALIGCGVISPQEDPKVFTACAGK110PLEL*LVSMNG 114

RESULT 13

US-09-615-192A-329
Sequence 329, Application US/09615192A
Patent No. 6410718
GENERAL INFORMATION:
APPLICANT: Bloksberg, Leonard N.
TITLE OF INVENTION: Materials and Methods for the
FILE REFERENCE: 11000.1003c40
CURRENT APPLICATION NUMBER: US/09/615.192A
CURRENT FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 08/975,316
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: US 08/713,000
PRIOR FILING DATE: 1996-09-11
PRIOR APPLICATION NUMBER: US 09/169,789
PRIOR FILING DATE: 1998-10-09
NUMBER OF SEQ ID NOS: 405
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 329
LENGTH: 97
TYPE: PR
ORGANISM: Eucalyptus grandis
US-09-615-192A-329

Query Match 10.7%: Score 389; DB 4; Length 97;
Best Local Similarity 76.3%: Pred. No. 9.3e-28;
Matches 74; Conservative 12; Mismatches 11; Indels 0; Gaps 0;

QY 182 NNNITPCLPLKITTASDPLVSLYAGLLTGPNSKAVGPTGEVLNKAFAAGVEGG 241
DB 1 NSGITPCLPLKISASGLVPPSYAGLLTGPNSKAVGPTGEVLNKAFAAGVEGG 60
QY 242 PFELOPKKGLAVNGVAGSCMAVLPDANVIALIS 278
DB 61 PFELOPKKGLAVNGVAGSCMAVLPDANVIALIS 97

RESULT 14

US-09-615-192A-331
Sequence 331, Application US/09615192A
Patent No. 6410718
GENERAL INFORMATION:
APPLICANT: Bloksberg, Leonard N.
APPLICANT: Havukkala, Ilkka
TITLE OF INVENTION: Materials and Methods for the
FILE REFERENCE: 11000.1003c40
CURRENT APPLICATION NUMBER: US/09/615.192A
CURRENT FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 08/975,316

PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: US 08/713,000
PRIOR FILING DATE: 1996-09-11
PRIOR APPLICATION NUMBER: US 09/169,789
PRIOR FILING DATE: 1998-10-09
NUMBER OF SEQ ID NOS: 405
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 331
LENGTH: 132
TYPE: PR
ORGANISM: Pinus radiata
US-09-615-192A-331

Query Match 7.6%: Score 276; DB 4; Length 132;
Best Local Similarity 58.5%: Pred. No. 2.7e-17;
Matches 55; Conservative 13; Mismatches 24; Indels 2; Gaps 1;

QY 16 IKDPLMVGAAALVNGSHLDEKVKVAFPRKPVVLCGETLTVSQVAGIAANSDIVKV 75
DB 41 IADPLMVGAAALVNGSHLDEKVKVAFPRKPVVLCGETLTVSQVAGIAANSDIVKV 98
QY 76 ELSEAARAGVAKASDWMVESNNKGTDSGVTTGF 109
DB 99 ELSEAARAGVAKASDWMVESNNKGTDSGVTTGF 132

RESULT 15

US-09-268-347-36
Sequence 36, Application US/09268347
Patent No. 6335182
GENERAL INFORMATION:
APPLICANT: Icosmure, Shucua M.
TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS
FILE REFERENCE: 1038-860
CURRENT APPLICATION NUMBER: US/09/268.347
CURRENT FILING DATE: 1999-03-16
NUMBER OF SEQ ID NOS: 54
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 36
LENGTH: 2411
TYPE: PR
ORGANISM: Haemophilus influenzae
US-09-268-347-36

Query Match 3.4%: Score 122.5; DB 4; Length 2411;
Best Local Similarity 20.7%: Pred. No. 0.34;
Matches 144; Conservative 81; Mismatches 247; Indels 223; Gaps 33;

QY 2 ENGNHNVGVNEICIKDPLMVGVA---AEALVGSHPDEKVKVAFPRKPVVLCGETL 57
DB 347 ESDGDTTASVT---KIDNGSITVYDAKVGDLKFDSDKKIVAD---TTALITV 395
QY 58 VSGVAGIAA---ANDSDIVKVELSEARAGVAKASDWMVESNNKGTDSGVTT 107
DB 396 GGVAVARIAKADHKKIVNAGDILVTAIGNISWAKAEADIPDCALEGISKQGEVKKACTV 455
QY 108 GFATSHRRTKQGA---LQKELIRPLNAGIF 136
DB 456 TFKAGRNILKVGQGANFTVSLDALTGLTSLITLVSTVWGNDAKIVIKKDLITTPNG 513
QY 137 GNCITSHLPHSATRAAM---IVRINTILQOYSGIRRELLDPAITKPLNNITPCL 189
DB 514 NGITGTNTI---SVTKDIGKAGNKAITVNASGLRAYDANDFDVNNSTADLNHVEDAYK 571
QY 190 ---PLRGITVAS---GDLVPLSYAGLLTGPNSKAVGPTGEVLNKAFAAGVEGG 241
DB 572 GLINLEKKNANOPVATDSTAATVGLKIKGVVSTKNGTKEHSQVYQADPVL---F 626
QY 234 AAGV---EGGFEL---QPKGIALV---KATA 258
DB 627 TGGAAATVTSKSENKHTITVSAETKALUSLEKGDITKIKLVNDQNDNLVTVNNKTA 686
QY 259 VGSMAVLPDANVIALISVIAISAFPAVVGCKPFTDHLITKIKHHPGQIFAALIMY 318

```

Db 687 VIKGGEIVTKGA-----:|||:| 730
QY 319 ILDSDYKA---AQVHEMDPLQFKUDRYALr---TSPWMLQPPJIEVIRSS-----365
Db 731 INSAATFVKTEI/TTSTIDEDNPTDNGKD--ALKAGDTLTPKAGKNLKVXKDKNITFDL 788
QY 366 TFMIEREINSVNDNPLIDVSRNKALHGNFQGTPIGVSMNTRLAIAIGK-LMFAQFS 424
Db 789 AKNLEVKTAKYSDTLT-----GKN---TPTGTTATPKYNTITSTADGLNFAKET- 835
QY 425 LVNDPYNNGLPNSISGGRNPGLDYGFQGEIIMASVCSSEIOLFANPVTHNVOSAFQHNOD 484
Db 836 -----ADASGSKNYVL---KG-----IATVITPEPSAGAKSSHYD 866
QY 485 VNSLGLISARKTAFAV---DILK---LMSSTYLVALQSIDLRLHEENMKSTVKNTVSQ 537
Db 867 LN---VDATKKSNAASIHDVYLRAGWNIQGNNGNVYVATVDYVNTIDSTIGTTVTVTO 922
QY 538 V-----AKKVLTMGVNGELHPSRCEKDI, 561
Db 923 KADGKADVKIIGAKTSVIKDHNGKL---FTGKDL 953

```

Search completed: March 29, 2003, 02:14:08
 Job time : 43 secs

GenCore version 5.1.4.P5.457A
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 29, 2003, 02:12:21 : Search time 182 Seconds

(without alignments)
229,410 Million cell updates/sec

Title: US-09-964-992A-1

Perfect score: 1644

1 MENGNHNVGVNELC IKDPL TIDPLETLYMNEPLPT 711

Sequence: 1

Scoring table: H10SUM62

Gapop 10.0, Gapext 0.5

Searched: 237916 seqs, 58723674 residues

Total number of hits satisfying chosen parameters: 237916

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 100%

Listing first 45 summaries

Database:

Published Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOM.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOM.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PTUS_PUBCOM.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOM.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOM.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOM.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOM.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3644	100.0	711	9	US-09-964-992A-1
2	3290.5	90.3	713	9	US-09-964-992A-2
3	3208	88.0	666	9	US-09-964-992A-5
4	2992	82.1	717	10	US-09-854-122-49
5	2728.5	74.9	700	10	US-09-854-122-50
6	1337	36.7	304	10	US-09-854-122-47
7	1332	46.6	303	10	US-09-854-122-48
8	1025	28.1	720	10	US-09-939-408A-13
9	994.5	27.3	716	10	US-09-765-873A-8
10	994.5	27.3	716	10	US-09-939-408A-19
11	991.5	27.2	716	10	US-09-765-873A-10
12	991.5	27.2	716	10	US-09-765-873A-34
13	990.5	27.2	716	10	US-09-765-873A-35
14	990.5	27.2	716	10	US-09-765-873A-37
15	988	27.1	686	10	US-09-765-873A-32
16	986.5	27.1	716	10	US-09-765-873A-38
17	984.5	27.0	716	10	US-09-765-873A-36
18	980.5	26.9	716	10	US-09-765-873A-33
19	971.5	26.7	713	10	US-09-939-408A-17

20	967	25.5	726	10	US-09-939-408A-21	Sequence 21, Appl
21	917	25.2	720	10	US-09-939-408A-30	Sequence 30, Appl
22	844.5	23.2	740	10	US-09-939-408A-15	Sequence 15, Appl
23	529	14.5	509	10	US-09-833-745-12085	Sequence 12085, A
24	527	14.5	513	10	US-09-833-745-50	Sequence 50, Appl
25	521.5	14.3	506	10	US-09-833-745-13960	Sequence 13960, A
26	518	14.2	513	10	US-09-833-745-51	Sequence 51, Appl
27	517	14.2	513	10	US-09-833-745-58	Sequence 58, Appl
28	513	14.1	509	10	US-09-833-745-34	Sequence 34, Appl
29	508	13.9	513	10	US-09-833-745-49	Sequence 49, Appl
30	505	13.9	508	10	US-09-833-745-40	Sequence 40, Appl
31	502	13.8	513	10	US-09-833-745-43	Sequence 43, Appl
32	501	13.7	513	10	US-09-833-745-53	Sequence 53, Appl
33	489	13.4	491	10	US-09-833-745-5671	Sequence 5671, Ap
34	489	13.4	504	10	US-09-833-745-2692	Sequence 2692, A
35	479	13.1	513	10	US-09-833-745-52	Sequence 52, Appl
36	476	13.1	504	10	US-09-833-745-95	Sequence 95, Appl
37	476	13.1	516	10	US-09-833-745-41	Sequence 41, Appl
38	473	13.0	513	10	US-09-833-745-46	Sequence 46, Appl
39	469.5	12.9	557	10	US-09-833-745-36	Sequence 36, Appl
40	468.5	12.9	513	10	US-09-833-745-92	Sequence 92, Appl
41	467.5	12.8	657	10	US-09-833-745-47	Sequence 47, Appl
42	461	12.7	513	10	US-09-833-745-48	Sequence 48, Appl
43	461	12.7	513	10	US-09-833-745-48	Sequence 48, Appl
44	458.5	12.6	513	10	US-09-833-745-45	Sequence 45, Appl
45	455.5	12.5	635	10	US-09-833-745-38	Sequence 38, Appl

ALIGNMENTS

RESULT 1
US-09-964-992A-1
Sequence 1, Application US/09964992A
Patent No. US20020173633A1
GENERAL INFORMATION:
APPLICANT: Sallveit, Mikal E.
APPLICANT: Campos, Reinaldo
APPLICANT: No. US20020173633A1logaki, Hiroyuki
APPLICANT: Ssalow, Trevor
TITLE OF INVENTION: The Regents of the University of California
TITLE OF INVENTION: Characterization of Phenylalanine Ammonia Lyase (PAL)
FILE REFERENCE: 023070-124500ms
CURRENT APPLICATION NUMBER: US/09/964,992A
PRIOR FILING DATE: 2002-07-09
PRIOR APPLICATION NUMBER: US 60/245,956
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 711
TYPE: PRT
ORGANISM: Lactuca sativa
FEATURE:
OTHER INFORMATION: Lettuce phenylalanine ammonia lyase (PAL)
US-09-964-992A-1

Query Match 100.0% Score 3644; DB 9; Length 711;
Post-local Similarity 100.0% Prev N. 1, Gap 272;
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MENGNHNVGVNELC IKDPLNMVAAATGSHIDVKKMAVFKKPVVKGITLTVS0 60
QY 1 MENGNHNVGVNELC IKDPLNMVAAATGSHIDVKKMAVFKKPVVKGITLTVS0 60
DB 1 MENGNHNVGVNELC IKDPLNMVAAATGSHIDVKKMAVFKKPVVKGITLTVS0 60
QY 61 VAGIAAANDSTYKVELSEAKAGVAKSSDWMESNNKIDTSYGVYIGCATSHRRK0G 120
QY 61 VAGIAAANDSTYKVELSEAKAGVAKSSDWMESNNKIDTSYGVYIGCATSHRRK0G 120
DB 61 VAGIAAANDSTYKVELSEAKAGVAKSSDWMESNNKIDTSYGVYIGCATSHRRK0G 120
QY 121 GALKKELIFPLNNAIPFNSTETSHITPHSATPAAMVPIINTLGVSGTREFLEIAIKF 180
|||||

Db	121	GAJOKELPFLUAGLFGNGTETHSHLDHSATRAAMIVRIINTLLGQSGIRELLEALITKE	180
Qy	181	LNNNTIPCLPLRGSTTTASSGLVPLSTYIAGLLTGRPSKAVGPTGEVLAANEKAFAAAGVEG	240
Db	181	LNNNTIPCLPLRGSTTTASSGLVPLSTYIAGLLTGRPSKAVGPTGEVLAANEKAFAAAGVEG	240
Qy	241	GFFELDPKGGALVNGTAVSGMASMYLFPANVALTALISVLSAIPAFWOCKPEFTTHT	300
Db	241	GFFELDPKGGALVNGTAVSGMASMYLFPANVALTALISVLSAIPAFWOCKPEFTTHT	300
Qy	301	HKLKHHPGJLEAAAMEYILLEGSDYVKAQAQVHEMDPLQPKUDRYALKTSPWLPJLE	360
Db	301	HKLKHHPGJLEAAAMEYILLEGSDYVKAQAQVHEMDPLQPKUDRYALKTSPWLPJLE	360
Qy	361	VIRSTTKMLEREINSVNPNPLIDVSRKKALHGGNFGGTFPIGVSMDNRILALAIQKLMFA	420
Db	361	VIRSTTKMLEREINSVNPNPLIDVSRKKALHGGNFGGTFPIGVSMDNRILALAIQKLMFA	420
Qy	421	QFSELVNDYFVNNGLPSSNLISGGRNPSLUDYGRKCHTMAASYCSFICPLANPVTNHVQSAQ	480
Db	421	QFSELVNDYFVNNGLPSSNLISGGRNPSLUDYGRKCHTMAASYCSFICPLANPVTNHVQSAQ	480
Qy	481	HNQDVSJLJLSAKRTAENVOLDLKLMSSTYLVMLQYSILULHLEEMKSTYKNVSYAK	540
Db	481	HNQDVSJLJLSAKRTAENVOLDLKLMSSTYLVMLQYSILULHLEEMKSTYKNVSYAK	540
Qy	541	KVLTMGVNGELHPSRFCEKEDLIRVDEDEYFAVIDVGSQTPYLMAKRLQVLDHIALNNG	600
Db	541	KVLTMGVNGELHPSRFCEKEDLIRVDEDEYFAVIDVGSQTPYLMAKRLQVLDHIALNNG	600
Qy	601	ETEKNTNTSIFOKIATFHEBELKVLLPKREVGVRIATYBNDTLSIPNKAQNSPYLRPVR	660
Db	601	ETEKNTNTSIFOKIATFHEBELKVLLPKREVGVRIATYBNDTLSIPNKAQNSPYLRPVR	660
Qy	661	EELGRGLTGEKVTSPGPEPFRVFTPAKCGGIIIDPLLECTGAGNCPPLPIC	711
Db	661	EELGRGLTGEKVTSPGPEPFRVFTPAKCGGIIIDPLLECTGAGNCPPLPIC	711
RESULT 2			
US-09-964-992A-2			
Sequence 2, Application US/09964992A			
Patent No. US2002017363A1			
GENERAL INFORMATION:			
APPLICANT: Salivetti, Mikel E			
APPLICANT: Campos, Reinaldo			
APPLICANT: No. US20020173633A10gaki, HiroYuki			
APPLICANT: Suslow, Trevor			
TITLE OF INVENTION: Characterization of Phenylalanine Ammonia-lyase (PAL)			
FILE OF INVENTION: Gene in Wounded Lettuce tissue			
FILE REFERENCE: 02307-124500ms			
CURRENT APPLICATION NUMBER: US/09/964, 992A			
CURRENT FILING DATE: 2002-07-09			
PRIOR APPLICATION NUMBR: US 60/235, 956			
PRIOR FILING DATE: 2000-09-26			
NUMBER OF SEQ ID NOS: 13			
SOFTWARE: PatentIn Ver. 2.1			
SEQ ID NO 2			
LENGTH: 713			
TYPE: PRT			
ORGANISM: Lactuca sativa			
FEATURE:			
OTHER INFORMATION: lettuce phenylalanine ammonia-lyase (PAL) 2			
OTHER INFORMATION: (USPAL2)			
US-09-964-992A-2			

```
Query Match:      90.3%; Score 3290.5; LH 9; Length 713;
Best Local Similarity 90.7%; Fred. No. 8.6e 245;
Matches 645; Conservative 29; Mismatches 32; Indels 5; Gaps 3;

OY    1  MENGNHNGVNNELGK-DPIIMGVAAEALDGSIHDEVKKVAEEFRKPYYVKLGSGFRTIVS 59
|| :| |: | :|: ||::||::|:| ||::||::||::||::||::||::||::||::||::
```

```

Db      6 MEVSHQNGRAECYKGDPLLMGMAABSLKSHLDEYKRMVAEPRKPVYRLGGETLTJVS 65
QY      60 QVASIAAANDSDYVKELSEARAGYKASDPMYMSMKGTDSYVTTGCAITSHRTKO 119
Db      66 QVAAL -AASNAAYKVELSETAAAGYKASDPMYMSMKGTDSYVTTGCAITSHRTKE 124
QY      120 GCAJOKELIPLFLAAGLIGWNLIEISHPLPSAUKAMIVKINILUQYSGIRFELLEITK 179
Db      125 GGALOKELIFLNAAGIFGNTESTHTPLPSATRAAMLVIRINTLLOGSGIRFELLEITK 184
QY      180 FLNNNTPLCLPLRGITIASGDVPLSYIAGLLTGRPNKAVPGTGEVLNAEKAPAAQVE 239
Db      185 FLNNNTVPLPLRGITIASGDVPLSYIAGLLTGRANSKAGPGEVLNAEKAPAAQYV 244
QY      240 GGFELQPKELILVNGTAVGSMASMYLFEDNVAALLSEVLSPFEDVMAQCAFEDHIL 299
Db      245 GGFELQPKELILVNGTAVGSMASMYLFEDNVAALLSEVLSPFEDVMAQCAFEDHIL 304
QY      300 IHLKLNHJ3J1EAAALMEY1LMDSDYVKAACVNHHPMDLQMKQJUKYALIKI SHQMDJ3QJ1 359
Db      305 THKLKHHPQIEAAALMEY1LMDSDYVKAACVNHHPMDLQPKODRALRTSPQJLSPQJ1 364
QY      360 EVTFSSTKMLIERELNSVNONPLLDVSRKNALHGSNFQSTPIGVSMDNTPLAIAIGKLM 419
Db      365 EVTFSSTKMLIERELNSVNONPLLDVSRKNALHGSNFQSTPIGVSMDNTPLAIAIGKLM 424
QY      420 AUFSELVNDYVNNGLSPYNSLUGRNPJLDYGFKGXSEIAAASCYSELOFLANPVINHVSAE 479
Db      425 AUFSELVNDYVNNGLSPYNSLUGRNPJLDYGFKGXSEIAAASCYSELOFLANPVINHVSAE 484
QY      480 OHNDVNSLGLISAKRTAEPAVDILKMSSTYVVALCOSIDILPHLEBNKSKIVKNTVSQVA 539
Db      485 OHNDVNSLGLISAKRTAEPAVDILKMSSTYVVALCOSIDILPHLEBNKSKIVKNTVSQVA 544
QY      540 KKVLTGCVNGELHPSRCECDLLRVYDREYVAYIDVQCSGTPLMOKLKOYIVDAHLN 599
Db      545 KKVLTGCVNGELHPSRCECDLLRVYDREYVAYIDVQCSATYPLMOKLKOYIVDAHLN 604
QY      600 GETEKNTNTSIFOKIATFEBELKVLPLKEVEGVRATYENDTLSPNRIKACRSYPIRYV 659
Db      605 ---ENDAGTSIFOKISFEFEELKAVLPKVEVEGVRASVESTLTITPNRIKCRSYPITRYV 661
QY      660 REELGRFLTGEKVYSPCEPEFDVFPAMKGOITDPLLELCJGKNNCPRI1 710
Db      662 REELGTGFLTGEVYSPGPEFDKVFALCKGHIIDPLECYGKNNGVPLT1 712

```

RESULT 3
 US-09-964-992A-5
 : Sequence: 5, Application US/09964992A
 : Patent No. US20020173633A1
 : GENERAL INFORMATION:
 : APPLICANT: Saltveit, Mikal E.
 : APPLICANT: Campos, Reinaldo
 : APPLICANT: No. US20020173633A1ogaki, Hiroyuki
 : APPLICANT: Suslow, Trevor
 : APPLICANT: The Regents of the University of California
 : TITLE OF INVENTION: Characterization of Phenylalanine
 : TITLE OF INVENTION: Gene in Wounded Lettuce Tissue
 : FILE REFERENCE: 023070-124500US
 : CURRENT APPLICATION NUMBER: US/09/964,992A
 : CURRENT FILING DATE: 2002-07-09
 : PRIOR APPLICATION NUMBER: US 60/735,956
 : PRIOR FILING DATE: 2000-09-26
 : NUMBER OF SEQ ID NOS: 13
 : SOFTWARE: PatentIn Ver. 2.1
 : SEQ ID NO 5
 : LENGTH: 666
 : TYPE: PRT
 : ORGANISM: Helianthus annuus
 : FEATURE:
 : OTHER INFORMATION: sunflower phenylalanine ammonia-lyase (PAL)
 : US-09-964-992A-5

Query Match 88.0% Score 3208 DB 9 Length 666;
 Host Local Similarity 94.6% Pred No 1 7c-238
 Matches 628; Conservative 12; Mismatches 24; Indels 0; Gaps 0.

QY 1 MGNHNGVYNELC IKDPLMGAVALTSGHLEDEYKKNVAEFRRPVYKLGSETLTVSQ 60
 DB 1 MGNHNGVYNELC IKDPLMGAVALTSGHLEDEYKKNVAEFRRPVYKLGSETLTVSQ 60
 QY 61 VAGIAAADSDTVKVEHSEARACVAKSSDWVMSNMKGTDSYGVTTGPGATSHRRTKOC 120
 DB 61 VAGIAAADSDTVKVEHSEARACVAKSSDWVMSNMKGTDSYGVTTGPGATSHRRTKOC 120
 QY 121 GALOKELREFINAGIFGNETSHLPHSATRAMIVRITLLQGYSGIFEELEATTKF 180
 DB 121 GALOKELREFINAGIFGNETSHLPHSATRAMIVRITLLQGYSGIFEELEATTKF 180
 QY 181 LNNITPLPLRGTTTASGDI VPLSYTAGLTPRPNKAVGPTGEVLNKAFAAGVEG 240
 DB 181 LNNITPLPLRGTTTASGDI VPLSYTAGLTPRPNKAVGPTGEVLNKAFAAGVEG 240
 QY 241 GPFELQPKGALVNGTAVGSGMASMVLFDANVALLSFVLSAIFAHVMOGKRPFTDHLT 300
 DB 241 GPFELQPKGALVNGTAVGSGMASMVLFDANVALLSFVLSAIFAHVMOGKRPFTDHLT 300
 QY 301 HKLKHHPQVLEAAALMEYLDSGVYKAAKVNHEMDLQPKYUKALRTSPQMLSPQLE 360
 DB 301 HKLKHHPQVLEAAALMEYLDSGVYKAAKVNHEMDLQPKYUKALRTSPQMLSPQLE 360
 QY 361 VRSSTKMEIEEINSVNDPLDVSRRKALHGNFUCSTPGVSKMONTRLAIAAGLMEFA 420
 DB 361 VRSSTKMEIEEINSVNDPLDVSRRKALHGNFUCSTPGVSKMONTRLAIAAGLMEFA 420
 QY 421 QSEELVNFYNNGLPSNLSGGRNPILDYCKGCEIAMAASVCELOFLANVTHVVSAGQ 480
 DB 421 QSEELVNFYNNGLPSNLSGGRNPILDYCKGCEIAMAASVCELOFLANVTHVVSAGQ 480
 QY 481 HNOVNSIAGLISAKRTAAVADILKLMSTYVALCOSIDLRHLEENKSTVTKTVSOVAK 540
 DB 481 HNOVNSIAGLISAKRTAAVADILKLMSTYVALCOSIDLRHLEENKSTVTKTVSOVAK 540
 QY 541 KVLTVNGELHPSRCEKDLRVADREYFAYIDVCSSTYPLMOKLRQVLDHALNNG 600
 DB 541 KVLTVNGELHPSRCEKDLRVADREYFAYIDVCSSTYPLMOKLRQVLDHALNNG 600
 QY 601 ETEKNTNTSIFOKIATPEEELKVLIPKVEGVRITAYENDTLSPNRIKACRSYPIVHFVR 660
 DB 601 ETEKNTNTSIFOKIATPEEELKVLIPKVEGVRITAYENDTLSPNRIKACRSYPIVHFVR 660
 QY 661 DELG 664
 DB 661 DELG 664
 RESULT 4
 US-09-854-122-49
 : Sequence 49, Application US/09854122
 : Patent No. US20020016980A1
 : GENERAL INFORMATION:
 : APPLICANT: ALBERTE, RANDALL S.
 : TITLE OF INVENTION: "TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
 : FILE REFERENCE: PHA-007.01
 : CURRENT FILING DATE: US/09/854,122
 : PRIOR APPLICATION NUMBER: 2001-09-10
 : PRIOR FILING DATE: 60/202,529
 : NUMBER OF SEQ ID NOS: 51
 : SOFTWARE: PatentIn Ver. 2.1
 : SEQ ID NO 49
 : TYPE: PRT
 : ORGANISM: Arabidopsis thaliana

US-09-854-122-49

Query Match 82.1% Score 2922 DB 19 Length 717;
 Host Local Similarity 81.2% Pred No 7 9c-222
 Matches 575; Conservative 52; Mismatches 79; Indels 2; Gaps 1;

QY 4 GNVHNGVYNELC IKDPLMGAVALTSGHLEDEYKKNVAEFRRPVYKLGSETLTVSQVAG 64
 DB 12 GNVHNGVYNELC IKDPLMGAVALTSGHLEDEYKKNVAEFRRPVYKLGSETLTVSQVAG 64
 QY 64 IAAANDSDTVKVEHSEARACVAKSSDWVMSNMKGTDSYGVTTGPGATSHRRTKOCAL 123
 DB 64 IAAANDSDTVKVEHSEARACVAKSSDWVMSNMKGTDSYGVTTGPGATSHRRTKOCAL 123
 QY 124 OKELIRPLINAGIFGNETSHLPHSATRAMIVRITLLQGYSGIFEELEATTKF 183
 DB 124 OKELIRPLINAGIFGNETSHLPHSATRAMIVRITLLQGYSGIFEELEATTKF 183
 QY 184 NITPCLPLRGTTTASGDI VPLSYTAGLTPRPNKAVGPTGEVLNKAFAAGVEG 243
 DB 184 NITPCLPLRGTTTASGDI VPLSYTAGLTPRPNKAVGPTGEVLNKAFAAGVEG 243
 QY 244 ELQPKGALVNGTAVGSGMASMVLFDANVALLSFVLSAIFAHVMOGKRPFTDHLT 303
 DB 244 ELQPKGALVNGTAVGSGMASMVLFDANVALLSFVLSAIFAHVMOGKRPFTDHLT 303
 QY 304 KHNPGQVLEAAALMEYLDSGVYKAAKVNHEMDLQPKYUKALRTSPQMLSPQLE 363
 DB 304 KHNPGQVLEAAALMEYLDSGVYKAAKVNHEMDLQPKYUKALRTSPQMLSPQLE 363
 QY 364 VRSSTKMEIEEINSVNDPLDVSRRKALHGNFUCSTPGVSKMONTRLAIAAGLMEFA 423
 DB 364 VRSSTKMEIEEINSVNDPLDVSRRKALHGNFUCSTPGVSKMONTRLAIAAGLMEFA 423
 QY 424 ELYNPFYNNGLPSNLSGGRNPILDYCKGCEIAMAASVCELOFLANVTHVVSAGQ 483
 DB 424 ELYNPFYNNGLPSNLSGGRNPILDYCKGCEIAMAASVCELOFLANVTHVVSAGQ 483
 QY 484 DVNSIAGLISAKRTAAVADILKLMSTYVALCOSIDLRHLEENKSTVTKTVSOVAK 543
 DB 484 DVNSIAGLISAKRTAAVADILKLMSTYVALCOSIDLRHLEENKSTVTKTVSOVAK 543
 QY 544 TMGVNGELHPSRCEKDLRVADREYFAYIDVCSSTYPLMOKLRQVLDHALNNG 603
 DB 544 TMGVNGELHPSRCEKDLRVADREYFAYIDVCSSTYPLMOKLRQVLDHALNNG 603
 QY 604 KNTNTSIFOKIATPEEELKVLIPKVEGVRITAYENDTLSPNRIKACRSYPIVHFVR 663
 DB 604 KNTNTSIFOKIATPEEELKVLIPKVEGVRITAYENDTLSPNRIKACRSYPIVHFVR 663
 QY 664 GGVLTGKRVTSQCHPEPDRVFAMCKQDILQPLAAGCQWNGEPLDPC 711
 DB 670 GGVLTGKRVTSQCHPEPDRVFAMCKQDILQPLAAGCQWNGEPLDPC 717
 RESULT 5
 US-09-854-122-50
 : Sequence 50, Application US/09854122
 : Patent No. US20020016980A1
 : GENERAL INFORMATION:
 : APPLICANT: ALBERTE, RANDALL S.
 : TITLE OF INVENTION: "TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
 : FILE REFERENCE: PHA-007.01
 : CURRENT FILING DATE: US/09/854,122
 : PRIOR APPLICATION NUMBER: 2001-09-10
 : PRIOR FILING DATE: 60/202,529
 : NUMBER OF SEQ ID NOS: 51
 : SOFTWARE: PatentIn Ver. 2.1
 : SEQ ID NO 50
 : TYPE: PRT

ORGANISM: Trillium aestivum
US-09-854-122-50

Query Match: 74.9%, Score 2728.5, DB 10, Length 700;
Best Local Similarity 75.9%; Pred. No. 1.4e-201;
Matches 527; Conservative 96; Mismatches 96; Indels 5; Gaps 3;

QY 18 DPLMGAVALGSHLDEYKKAWEFRKPVYKLTGSETLTVQVAGIAAANDSTIVKEL 77
DB 12 DPLMGAVALGSHLDEYKKAWEFRKPVYKLTGSETLTVQVAGIAAANDSTIVKEL 77
QY 78 SEARAGVAKASVWVMSMKNKGTDSYGVTTGEGATSHRRTKQAGALOKELIFELNAGIFG 137
DB 68 DESARAGVAKASVWVMSMKNKGTDSYGVTTGEGATSHRRTKQAGALOKELIFELNAGIFG 127
QY 138 NGFETSHLPHSANTRAMIVRINTLLOQSGIFELFELMIFKLNNTIPCLPUGCTTAA 197
DB 128 TGTD-GHVLPAATRAMLKVNTLLQGSGLIFELFELMIFKLNNTIPCLPUGCTTAA 186
QY 198 SGDLVPLSTYAGLTLGRPNKAVGPTGEVLENAKFAAGVSGFELPKBGLAVNGT 257
DB 187 SGDLVPLSTYAGLTLGRPNKAVGPTGEVLENAKFAAGVSGFELPKBGLAVNGT 246
QY 258 AVSGMASVPLEDANVTALLSEVLSAIFAEVNOGKPEFTDHLTHKLHHPGQIEAAIAME 317
DB 247 AVSGMASVPLEDANVTALLSEVLSAIFAEVNOGKPEFTDHLTHKLHHPGQIEAAIAME 306
QY 318 YLLDGSYVKAQKVHMDPQKPKODRYALRTSPQWLGQIEVIRHSTKMLEREINSYN 377
DB 307 HILEGSSYMLAKKGLGELPDKMPKODRYALRTSPQWLGQIEVIRHSTKMLEREINSYN 366
QY 378 DNPLIDVSRKALHGNFQGTPIGVSMNTRLAIAIGLMEFAQSELVNDFYNGGLPSN 437
DB 367 DNPLIDVSRKALHGNFQGTPIGVSMNTRLAIAIGLMEFAQSELVNDFYNGGLPSN 426
QY 438 LSGGRNSLDYGRKGGEIAVASCSELOFLANPYTNHVOASBOHNDVNSLISARKTA 497
DB 427 LSGGRNSLDYGRKGGEIAVASCSELOFLANPYTNHVOASBOHNDVNSLISARKTA 486
QY 498 EAVDILKMSSTYVALCOSIDIRHLENNKSTVKNFVSQVAKKVTMCGNGLHPSRPG 557
DB 487 EAVDILKMSSTYVALCOSIDIRHLENNKSTVKNFVSQVAKKVTMCGNGLHPSRPG 546
QY 558 EKDLIRVYDREYFAYIDVCSSTYPLMKRLROYLVDAHLNNEFKNTNTSIFOKIATF 617
DB 547 EKDLIRVYDREYFAYIDVCSSTYPLMKRLROYLVDAHLNNEFKNTNTSIFOKIATF 606
QY 618 EELKVLLEPEVEGVRIAYENDLSIPNRIKARSTYPLRFVEELRGFELTGEVTSFG 677
DB 607 EELKVLLEPEVEGVRIAYENDLSIPNRIKARSTYPLRFVEELRGFELTGEVTSFG 666
QY 678 EEDRVATAMCKCOIIDPLKLCGNNGPPLPIC 711
DB 667 EEDRVATAMCKCOIIDPLKLCGNNGPPLPIC 700

RESULT 6
US-09-854-122-47

Sequence 47, Application US/09854122
Patent No. US20020016980A1
GENERAL INFORMATION:
APPLICANT: ALBERTE, RANDALL S.
TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
FILE REFERENCE: PHA-007.01
CURRENT APPLICATION NUMBER: US/09/854,122
CURRENT FILING DATE: 2001-09-10
PRIOR APPLICATION NUMBER: 60/202,529
PRIOR FILING DATE: 2000-05-10
NUMBER OF SEQ ID NOS: 51
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 47
LENGTH: 304

TYPE: PRT
ORGANISM: Zostera marina
US-09-854-122-47

Query Match: 36.7%, Score 1337, DB 10, Length 304;
Best Local Similarity 87.7%; Pred. No. 4.1e-95;
Matches 256; Conservative 20; Mismatches 16; Indels 0; Gaps 0;

QY 339 QKPKODRYALRTSPQWLGQIEVIRHSTKMLEREINSVNDNLIDVSRKALHGNFQGT 398
DB 13 KSKODRYALRTSPQWLGQIEVIRHSTKMLEREINSVNDNLIDVSRKALHGNFQGT 72
QY 399 PIGVSMNTRFLAIAAIGKLMFAQFSELVNDFFYNNGLPSNLSCGRNPSTDYGRKGEIAMA 458
DB 73 PIGVSMNTRFLAIAAIGKLMFAQFSELVNDFFYNNGLPSNLSCGRNPSTDYGRKGEIAMA 132
QY 459 SYCSELOFLANPYTNHVOASBOHNDVNSLGLISARKTAEAVDILKMSSTYVALCOSI 518
DB 133 SYCSELOFLANPYTNHVOASBOHNDVNSLGLISARKTAEAVDILKMSSTYVALCOSI 192
QY 519 DLKHEENLKASVKNIVSOVAKKVLTMGVNGGLHPSRCEKDLKVVDRYVVFSTIDDPG 578
DB 193 DLKHEENLKASVKNIVSOVAKKVLTMGVNGGLHPSRCEKDLKVVDRYVVFSTIDDPG 252
QY 579 SGTYPLMOKLROYLVDAHLNNGEFTKNTSIFOKIATFEEELKVLLEPREVE 630
DB 253 SATYPLMOKLROYLVDAHLNNGEFTKNTSIFOKIATFEEELKVLLEPREVE 304

RESULT 7
US-09-854-122-48

Sequence 48, Application US/09854122
Patent No. US20020016980A1
GENERAL INFORMATION:
APPLICANT: ALBERTE, RANDALL S.
TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
FILE REFERENCE: PHA-007.01
CURRENT APPLICATION NUMBER: US/09/854,122
CURRENT FILING DATE: 2001-09-10
PRIOR APPLICATION NUMBER: 60/202,529
PRIOR FILING DATE: 2000-05-10
NUMBER OF SEQ ID NOS: 51
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 48
LENGTH: 303
TYPE: PRT
ORGANISM: Zostera marina
US-09-854-122-48

Query Match: 36.6%, Score 1332, DB 10, Length 303;
Best Local Similarity 87.6%; Pred. No. 9.9e-95;
Matches 255; Conservative 20; Mismatches 16; Indels 0; Gaps 0;

QY 339 QKPKODRYALRTSPQWLGQIEVIRHSTKMLEREINSVNDNLIDVSRKALHGNFQGT 398
DB 13 KSKODRYALRTSPQWLGQIEVIRHSTKMLEREINSVNDNLIDVSRKALHGNFQGT 72
QY 399 PIGVSMNTRFLAIAAIGKLMFAQFSELVNDFFYNNGLPSNLSCGRNPSTDYGRKGEIAMA 458
DB 73 PIGVSMNTRFLAIAAIGKLMFAQFSELVNDFFYNNGLPSNLSCGRNPSTDYGRKGEIAMA 132
QY 459 SYCSELOFLANPYTNHVOASBOHNDVNSLGLISARKTAEAVDILKMSSTYVALCOSI 518
DB 133 SYCSELOFLANPYTNHVOASBOHNDVNSLGLISARKTAEAVDILKMSSTYVALCOSI 192
QY 519 DLKHEENLKASVKNIVSOVAKKVLTMGVNGGLHPSRCEKDLKVVDRYVVFSTIDDPG 578
DB 193 DLKHEENLKASVKNIVSOVAKKVLTMGVNGGLHPSRCEKDLKVVDRYVVFSTIDDPG 252
QY 579 SGTYPLMOKLROYLVDAHLNNGEFTKNTSIFOKIATFEEELKVLLEPREVE 630
DB 253 SATYPLMOKLROYLVDAHLNNGEFTKNTSIFOKIATFEEELKVLLEPREVE 303

RESULT 8

US-09-939-408A-13
Sequence 13, Application US/09/939408A
Patent No. US20020102712A1
GENERAL INFORMATION:
APPLICANT: Yoshida, Nobeta
APPLICANT: Koolstra, Anna
TITLE OF INVENTION: Phenylalanine Ammonia Lyase Polypeptide and
TITLE OF INVENTION: Polynucleotide Sequences and Methods of Obtaining and
TITLE OF INVENTION: Using Same
FILE REFERENCE: 29479/500NSCA
CURRENT APPLICATION NUMBER: US/09/939,408A
PRIOR FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 09/624,693
PRIOR FILING DATE: 2000-07-24
PRIOR APPLICATION NUMBER: PCT/US01/23270
PRIOR FILING DATE: 2001-07-24
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Patent Vei. 2.0
SEQ ID NO 13
LENGTH: 720
TYPE: PRT
ORGANISM: Rhodotorula glutinis
FEATURE:
NAME/KEY: unsure
LOCATION: (153)
OTHER INFORMATION: Other Information: Xaa - Val or Ala
US-09-939-408A-13

Query Match 28.18, Score 1025, DA 10, Length 720,
Best Local Similarity 47.98, Pred. No. 1.5e-70,
Matches 267, Conservative 127, Mismatches 259, Indels 52, Gaps 18.

QY 23 CVAALALGSHLDIVKKVAVPKPKPVKIGCTITVSOVAGIAANDSTVKKELSHAR 82
DB 40 GLCHAAHQSLIEVLELSDPDIIDVVELSDYSLIKDYVG-AAKGRKVRQNDDEL 97
QY 83 AGVAKSSWVMSNMKGTDSYVTTGFGATSHRTKGGALOKELIFELNAGI----- 135
DB 98 ARVAKSVLPKAGVQNSV-YGVITGCGSADPTKIELAVSLKALIRHQVGVPIPSXSS 155
QY 136 FONGTSHITLPHSATKAMIVKINTLQGYSGIRPELEATIKYNNITPCLPQGT 195
DB 156 FSVGPLENTLPLEVPRGAMVPIVNSLPGHSAYLVLELTFLNHPITPIVLRGSI 215
QY 196 TASCGLVPLSYAGLITGRPNRKA---VGPTEVLNAKKAPAAAGVCGPPELQPKHGIA 252
DB 216 SASGLPLSLYAGLITGRPNRKA---VGPTEVLNAKKAPAAAGVCGPPELQPKHGIA 273
QY 253 LVNGTAVGSGMAVLEFANVLLSEVLSAIEFVAGKRP-EPTIDILTKLKHPRQTE 311
DB 274 LVNGTAVGSGMAVLEFANVLLSEVLSAIEFVAGKRP-EPTIDILTKLKHPRQTE 333
QY 312 AAALIMETLLDSDY-VKAAQKVHMDPLQKQUDVALRTSPMLGPOIVIRSSTKMTE 370
DB 334 VARNIRLLSSGSAVEVEEYKVKDDEGLRQDRPLPSPQGLPVEDMMHVAITLS 393
QY 371 REINSVNDPLIDVSRNKALHCGNPGCTPGVSMNDTRILAAIGKIMFOPSHLVNDFY 430
DB 394 LEFNTTTPNPLIDVSRNKALHCGNPGCTPGVSMNDTRILAAIGKIMFOPSHLVNDFY 453
QY 431 NNGTSPNSGSRNPSLLYVFKKAEIEMASVYSELQFLANPTNIVVSAPVHNDVNSL 490
DB 454 NNGTSPNSGSRNPSLLYVFKKAEIEMASVYSELQFLANPTNIVVSAPVHNDVNSL 512
QY 491 ISAKRTAAVNDILKIMSTYVALQSDIDILHLENNKSTVKNVQSAKAVITMG--VN 548
DB 513 ISAKRTAAVNDILKIMSTYVALQSDIDILHLENNKSTVKNVQSAKAVITMG--VN 572
QY 549 G-ELHPSRFRCKDLKLVKREY-----VFAYLIDVSCITVPIVMOUKLVIVYHAIANGE 601
DB 572 G-ELHPSRFRCKDLKLVKREY-----VFAYLIDVSCITVPIVMOUKLVIVYHAIANGE 601

DB 573 ALALPVKKALKRLKQITTYDLEPRMHAIFY-----ATGT-----VVELSSSP 617

QY 602 TEKNTNTSI-FOKIATEEELKVLIPKEVESVRIAVENDTLSPNLIKAKSYPLREVR 660

DB 618 SANVTITAVNAMKVASAKATS--LTREVPN-FEMJIPSSGABAHVLSPTVLSFVR 674

QY 661 FELG-----RG-FLIGEKVTSPEGEFDPVFTAMCKGQIDPLILCL 700

DB 675 BELGQVAPRGVFGVQDETIGSNVSRITYCAIKDGRINIVLVKML 719

RESULT 9

US-09-765-873A-8
Sequence 8, Application US/09765873A
Patent No. US20010053847A1
GENERAL INFORMATION:
APPLICANT: Tang, Xiao-Song
TITLE OF INVENTION: HOPRODUCTION OF HARA-HYDROXYCINNAMIC ACID
FILE REFERENCE: PCT009 US CIP
CURRENT APPLICATION NUMBER: US/09/765,873A
PRIOR FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: US 99/627,216
PRIOR FILING DATE: 2000-07-27
PRIOR APPLICATION NUMBER: US 60/147,719
PRIOR FILING DATE: 1999-08-06
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Microsoft Office 97
SEQ ID NO 8
LENGTH: 716
TYPE: PRT
ORGANISM: Rhodotorula glutinis
US-09-765-873A-8

Query Match 27.38, Score 994.5, DA 10, Length 716,
Best Local Similarity 35.98, Pred. No. 3.2e-68,
Matches 270, Conservative 123, Mismatches 250, Indels 109, Gaps 23.

QY 8 NGVNEELCIKIDPLMMGYAFAALGSHL-----DEVKKVAAPKPKPVKIGCTITVS 59
DB 14 NGVNS---AKAVNAGASNTNLAVASHLPTVTVLVLEKMLAIPDSTLELGYSLNLG 70
QY 60 QVAGIAAANSTVKKVLESPAARAGVAKSSWVMSNMKGTDSYVTTGFGATSHRTK 119
DB 71 IV-VSAARKGPRPVKVSDEIRSKIPKSVFLPSQLSMV-YGVITGCGSADPTK 126
QY 120 GSALQKELIFELNAGI-----FNGTETSHITLPHSATKAMIVKINTLQGYSGIR 170
DB 127 ALSIDKALLHQLDGVLPSSPDSRLDQGLNS--IDPLVVRGAMITKVNSTIRGSAVR 184
QY 171 FELLRAITKPLNNITPQIPKCTITFASGDIPLSYAGLITGRPNRKA-VGPTG--EVL 227
DB 185 LVLEALTNFNLHGITPIVPLRGITISAGDLSPLSYIAAISGIPSKVIVHEGEKEL 244
QY 228 NAKKRAAACYVCGCPPELQPKRGIALVNGTAVGSGMAVLEFANVLLSEVLSAIEFV 287
DB 245 YAREFMALEFNLDP--VVIQPKREGGLVNGIAVSAAMATIALIHAMLSILSSITAMTYE 302
QY 288 VMGQKP-EFTDILTKLKHPRQTEAAAIMEDYLLDSDY-VKAAQKVHMDPLQKQUD 345
DB 303 AMVGAQSGFHPITLDVTRPHPTQIVKAGNIRKLKJGSRPAVNHREKVKVDDGCLIKQR 362
QY 346 YALRTSPMLGPOIVIRSSTKMIREINSVNDPLIDVSRNKALHCGNPGCTPGVSM 404
DB 363 YPLTPSPQMLPISVNDLITHAHVLITEGGSTDPNPLIEVNTSHHGCFVAAAVANFM 422
QY 405 DNTRIALAIKLMFAQVSHLVNDFYNNGLPSNLSGGRNLSIDYKCGKGLIAAASGSH 464
DB 423 KTRIGIAQIKINFQITTEMLNAGMNRGLPSCLA-AHDPSLSTRCKGLDIAAATSTEL 481
QY 465 QFLANPTNIVVSAPVHNDVNSLISAKRTAAVNDILKIMSTYVALQSDIDILHLENN 524
DB 482 GHLANPTNIVVSAPVHNDVNSLISAKRTAAVNDILKIMSTYVALQSDIDILHLENN 541

```
QY 525 ENNKSTVKNVTSQVAKKVLTMGVNSELHPSRCEKDLRVVDREVFAVIDVCSGTPL 584
DB 542 FEKK-----OFPA-----IVSLIDQH-----GSAMT 565
QY 585 MOKIROVLDHALNNGETEKNTSIFOKIATFEELKVLPEVEGV-----632
DB 566 GSNLRDELVE-KVKKTLAKRLQETNSYDLPVPMHAFSPAQTVVEVLSLSLAAYNA 624
QY 633 -RIAYENDTLSTIPNRK-----ACRSYP-----LYRPREELG-----RG--FL 668
DB 625 WKVAAAESALSTRQVRETFWSAASSTSPALSTLSPRTQILYAFVREELGVAKARGDVF 684
QY 669 TGEKVTSPGEEFDRVFTAMCKGQIIDPLECL 700
DB 685 GKQEVTT-GSNVSKTYEAIKSGRINVLKML 715

RESULT 10
US-09-939-408a-19
: Sequence 19, Application US/09939408A
: Patent No. US20020102712A1
: GENERAL INFORMATION:
: APPLICANT: Yoshida, Roberta
: APPLICANT: Koetsira, Anna
: TITLE OF INVENTION: Phenylalanine Ammonia Lyase Polypeptide and
: TITLE OF INVENTION: Polynucleotide Sequences and Methods of Obtaining and
: FILE REFERENCE: 29479/500NSCA
: CURRENT APPLICATION NUMBER: US/09/939,408A
: PRIOR FILING DATE: 2000-07-24
: PRIOR APPLICATION NUMBER: PCT/US01/23270
: PRIOR FILING DATE: 2001-07-24
: NUMBER OF SEQ ID NOS: 30
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 19
: LENGTH: 716
: TYPE: PRP
: ORGANISM: Rhodotorula toruloides
US-09-939-408a-19

Query Match 27.3%: Score 994.5; DB 10; Length 716;
Best Local Similarity 35.9%: Pred No 3.2e-68;
Matches 270; Conservative 123; Mismatches 250; Indels 109; Gaps 23;

QY 8 NGVYNELCIKDPLNMGVAAEALTGSHL-----DEVKKMVAEPRKPVVKGHTLTVS 59
DB 14 NGVAS---AKQAVNGASTNLAVAGSHLPTTYQYVDYVEKMLAAPDSTLELDGYSLMG 70
QY 60 QVAGIAAANDSDTVKVELSEAAAGYKASSDWMSMKNKGTDSYCTTGTCATSHKRTKQ 119
DB 71 DV--VSAARKGRPVVKDDEIRSKIDKSEVPIRSQLMSV--YGVPTGCHGSAADRTED 126
QY 120 GGLAKLELREFLNAGI-----FGNGTETSHTLPHSATRAAMIVRINTLGGYSGIR 170
DB 127 AISLQKALLEHOLCGVLPSSPFRIGRLGLEN--LPLFVVGKANTIKVNSLTRGHSAYR 184
QY 171 FEILEATIKFLNNNTTPCLPRTGTTASGDLVPLSYIAGLLTGRPNKA-VGPTG--EVL 227
DB 185 LVLEALTNELNIGITPIVPLRGTTISASGDSLPLSYIAAISGHPDSKVHVHNESEKIL 244
QY 228 NAEKAFVAAAGVEGFELOPKRGCLALVNGTAVSGMASMVLPDANVALLSVLSAIFAE 287
DB 245 YAREAAALFNLBP--VVLGPKRGIGLVNGTAVSASMATLALHDHMLSTLSQSLTAMTVE 302
QY 288 VMGKRP-EFTDHLTKLKHHPQIEAAAIMEXILDGSDY-VKAAQKVHEMDPLQPKQDR 345
DB 303 AMVGHASGFRPLADYTRKPRPTQIEVAGNIKKLJCSRRVAHVHEEVKVKDDHGLRDR 362
QY 346 YALRTSPQWLGPOI-EVYKSTKMIEREINSVNDNPLDIVSRNKALHCGNPGTPIGVSM 404
DB 363 YPLRTSPQWLGPLVSDLIHAHVALITEAGOSTDNDPLIDVENKTSHHGNGFQAAVANTM 422
```

```
QY 405 DNTRLAIAIKLMPAOPSELVNDVFYNNGLPSNLGSPNPSIDYGGKGETAMASCSSEL 464
DB 423 EKTRGLIAOIGKLNFTQLEMLNAGNNRQLPSCLA-AEPPSLYHCKGLDIAAAVYSEL 481
QY 465 QPLANPVNTHVQSAEHDNDVNSGLISARKTAEAVDILKMSSTYVALCOSIDRLHLE 524
DB 482 GHILANPVTHVCPAEAMANOAVNSLALISARPTESNDVLSLLATHLYCVALDILRAIE 541
QY 525 ENNKSTVKNVTSQVAKKVLTMGVNSELHPSRCEKDLRVVDREVFAVIDVCSGTPL 584
DB 542 FEKK-----OFPA-----IVSLIDQH-----GSAMT 565
QY 585 MOKIROVLDHALNNGETEKNTSIFOKIATFEELKVLPEVEGV-----632
DB 566 GSNLRDELVE-KVKKTLAKRLQETNSYDLPVPMHAFSPAQTVVEVLSLSLAAYNA 624
QY 633 -RIAYENDTLSTIPNRK-----ACRSYP-----LYRPREELG-----RG--FL 668
DB 625 WKVAAAESALSTRQVRETFWSAASSTSPALSTLSPRTQILYAFVREELGVAKARGDVF 684
QY 669 TGEKVTSPGEEFDRVFTAMCKGQIIDPLECL 700
DB 685 GKQEVTT-GSNVSKTYEAIKSGRINVLKML 715

RESULT 11
US-09-765-873A-10
: Sequence 10, Application US/09765873A
: Patent No. US20010053847A1
: GENERAL INFORMATION:
: APPLICANT: Tang, Xiao-Song
: TITLE OF INVENTION: BIOPRODUCTION OF PAKA-HYDROXYCINNAMIC ACID
: FILE REFERENCE: BC1009 US CIP
: CURRENT APPLICATION NUMBER: US/09/765,873A
: PRIOR FILING DATE: 2001-01-19
: PRIOR APPLICATION NUMBER: US 09/627,216
: PRIOR FILING DATE: 2000-07-27
: PRIOR APPLICATION NUMBER: US 60/147,719
: PRIOR FILING DATE: 1999-08-06
: NUMBER OF SEQ ID NOS: 38
: SOFTWARE: Microsoft Office 97
: SEQ ID NO 10
: LENGTH: 716
: TYPE: PRP
: ORGANISM: mutant from Rhodotorula glutinis
US-09-765-873A-10

Query Match 27.2%: Score 991.5; DB 10; Length 716;
Best Local Similarity 35.9%: Pred No 5.5e-68;
Matches 270; Conservative 122; Mismatches 251; Indels 109; Gaps 23;

QY 8 NGVYNELCIKDPLNMGVAAEALTGSHL-----DEVKKMVAEPRKPVVKGHTLTVS 59
DB 14 NGVAS---AKQAVNGASTNLAVAGSHLPTTYQYVDYVEKMLAAPDSTLELDGYSLMG 70
QY 60 QVAGIAAANDSDTVKVELSEAAAGYKASSDWMSMKNKGTDSYCTTGTCATSHKRTKQ 119
DB 71 DV--VSAARKGRPVVKDDEIRSKIDKSEVPIRSQLMSV--YGVPTGCHGSAADRTED 126
QY 120 GGLAKLELREFLNAGI-----FGNGTETSHTLPHSATRAAMIVRINTLGGYSGIR 170
DB 127 AISLQKALLEHOLCGVLPSSPFRIGRLGLEN--LPLFVVGKANTIKVNSLTRGHSAYR 184
QY 171 FEILEATIKFLNNNTTPCLPRTGTTASGDLVPLSYIAGLLTGRPNKA-VGPTG--EVL 227
DB 185 LVLEALTNELNIGITPIVPLRGTTISASGDSLPLSYIAAISGHPDSKVHVHNESEKIL 244
QY 228 NAEKAFVAAAGVEGFELOPKRGCLALVNGTAVSGMASMVLPDANVALLSVLSAIFAE 287
DB 245 YAREAAALFNLBP--VVLGPKRGIGLVNGTAVSASMATLALHDHMLSTLSQSLTAMTVE 302
QY 288 VMGKRP-EFTDHLTKLKHHPQIEAAAIMEXILDGSDY-VKAAQKVHEMDPLQPKQDR 345
```


LOCATION: (540) (540)
OTHER INFORMATION: X= Thr, Ala, Ser, Pro, Gly
US-09-765-873A-35

Query Match 27.2% Score 990.5; DB 10; Length 716;
Best Local Similarity 35.9%; Pred No. 6, 6e-68;
Matches 270; Conservative 122; Mismatches 251; Indels 109; Gaps 23;

QY 8 NGVNEICIKPPLMKGVAEALTSGL-----DEYKKVAEFKKYVKLGSETLTVS 59
DB 14 NGVAS---AKOAVNGASTNINLAAGSHLPTTOYVQVDIVKMLAAPDSTLELDGYSLNIG 70
QY 60 GYAGIAANDSDTYKVELSEARAGVAKASDPMVMSNKKGIDSGYVTTGCAISHHRTKO 119
DB 71 DY--VSAARKGRPRVAKDSDEIRSKIDKSVFELRSQLSMSV--YGVTTGGASDRTED 126
QY 120 GGALOKELIRPLMGI-----FCNGTSHLPHSATRAMIVRINTLIQYSGIR 170
DB 127 AISTLOKALLEHQLCGVLPSSFDSPFLRGLENS--LPLEVVRGAMTIRVNSLTRGHSVR 184
QY 171 FELLEAIKFLNNNTPTPLRGTTASGDLVPLSYAGLITGRPNKA--VGFTG--EVL 227
DB 185 LVLLEALNPLNHTTPTPLRGTTASGDLSPSTYIAAISCGRPSKVVHVEGKEKIL 244
QY 228 NAEKAPAAAGVEGFEFLQPEGLALNGTAVSGMASMVLFDANVLLSEVLALFAE 287
DB 245 VAREMALFNLER--VVLGPRKGGLVNGTAVSASMATLALHDAHMLSLSSLTAMTVE 302
QY 288 VMQGRP--EFTDHLTKLKHHPQGLEAAALMEYILDSGY--VKAQVKHEMDPIQKPRDR 345
DB 303 AMVAGAGSHFPLHDVTRPHPTQIEVAGNIRKILEGSEFAVHHEEVKVKVDDGILRODR 362
QY 346 VALRTPQWMLGPOI--EVRISSTKMIEREINSVNDPILIVSRKALHSGNFQSTPIGVS 404
DB 363 YTLRTPQWMLGPOI--EVRISSTKMIEREINSVNDPILIVSRKALHSGNFQSTPIGVS 422
QY 405 DNTLRLAIAIKLMAPOSELVNDPYNNGILPSNLSSGRNPFLDGRKSGELIAMAASYSEL 464
DB 423 EKTRLGLAOIGKLNFTQLEMLNAGMNRGLPSCLA--AEDESLSHCKGDLIAAAAYTSEL 481
QY 465 OFLANPVNHHQSAOHQOVNSLGLISARKTAAVLDLILMSTYVVALCOSIDLRIHLE 524
DB 482 GHLANPVTHVQAPMAQAVNSLALLISARKTESNVLSLLATHLHYCVLQALIDIRAXE 541
QY 525 ENMKSTVKNVTSQVAKKVLTMGVNCELHPSRFCEKDLIRVVDREYVAYIDVCSGYPL 564
DB 542 FEFKK-----QFGPA-----IVSLIDQH-----GSAMT 565
QY 585 MOKLRVLDHALNNGETEKNTNIFQKIATFEELKVLILPKREYGV-----632
DB 566 GSNLDELIV--KVNKTIAKRLKQTNSTYDLVRMHDAFSFAGTYVEVLSTSLAAVNA 624
QY 633 -RIAYENDTLISPNRIK-----ACRSYP-----LYRVRRELG---RG--FL 668
DB 625 WKVAAEASALISLTROVRETFMSASTSPALSYLSPRTQILVAVREBELGVAKRGDVL 684
QY 669 TGEKVTSPGEEDRVFTAMCKGQIIDLLECL 700
DB 685 GKQEVYI--GSNVSKITYEAIKSGRIANNVLLKM 715

RESULT 14

US-09-765-873A-37
Sequence 37, Application US/09765873A
Patent No. US20010053847A1
GENERAL INFORMATION:
APPLICANT: Tang, Xiao-Song
TITLE OF INVENTION: BIOPRODUCTION OF PARA-HYDROXYCINNAMIC ACID
FILE REFERENCE: H01009 US CIP
CURRENT APPLICATION NUMBER: US/09/765, 873A
CURRENT FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: US 09/627, 216
PRIOR FILING DATE: 2000-07-27

PRIOR APPLICATION NUMBER: US 60/147,719
PRIOR FILING DATE: 1999-08-06
NUMBER OF SEQ. ID NOS: 38
SOFTWARE: Microsoft Office 97
SEQ. ID NO 37
LENGTH: 716
TYPE: PRT
ORGANISM: mutant from Rhodotorula glutinis
FEATURE:
NAME/KEY: UNSURE
LOCATION: (149)..(149)
OTHER INFORMATION: X= Pro, Ala, Ser, Thr, Gly
NAME/KEY: UNSURE
LOCATION: (202)..(202)
OTHER INFORMATION: X= Val, Met, Leu, Cys
NAME/KEY: UNSURE
LOCATION: (540)..(540)
OTHER INFORMATION: X= Thr, Ala, Ser, Pro, Gly
US-09-765-873A-37

Query Match 27.2% Score 990.5; DB 10; Length 716;
Best Local Similarity 35.9%; Pred No. 6, 6e-68;
Matches 270; Conservative 122; Mismatches 251; Indels 109; Gaps 23;

QY 8 NGVNEICIKPPLMKGVAEALTSGL-----DEYKKVAEFKKYVKLGSETLTVS 59
DB 14 NGVAS---AKOAVNGASTNINLAAGSHLPTTOYVQVDIVKMLAAPDSTLELDGYSLNIG 70
QY 60 GYAGIAANDSDTYKVELSEARAGVAKASDPMVMSNKKGIDSGYVTTGCAISHHRTKO 119
DB 71 DY--VSAARKGRPRVAKDSDEIRSKIDKSVFELRSQLSMSV--YGVTTGGASDRTED 126
QY 120 GGALOKELIRPLMGI-----FCNGTSHLPHSATRAMIVRINTLIQYSGIR 170
DB 127 AISTLOKALLEHQLCGVLPSSFDSPFLRGLENS--LPLEVVRGAMTIRVNSLTRGHSVR 184
QY 171 FELLEAIKFLNNNTPTPLRGTTASGDLVPLSYAGLITGRPNKA--VGFTG--EVL 227
DB 185 LVLLEALNPLNHTTPTPLRGTTASGDLSPSTYIAAISCGRPSKVVHVEGKEKIL 244
QY 228 NAEKAPAAAGVEGFEFLQPEGLALNGTAVSGMASMVLFDANVLLSEVLALFAE 287
DB 245 VAREMALFNLER--VVLGPRKGGLVNGTAVSASMATLALHDAHMLSLSSLTAMTVE 302
QY 288 VMQGRP--EFTDHLTKLKHHPQGLEAAALMEYILDSGY--VKAQVKHEMDPIQKPRDR 345
DB 303 AMVAGAGSHFPLHDVTRPHPTQIEVAGNIRKILEGSEFAVHHEEVKVKVDDGILRODR 362
QY 346 VALRTPQWMLGPOI--EVRISSTKMIEREINSVNDPILIVSRKALHSGNFQSTPIGVS 404
DB 363 YTLRTPQWMLGPOI--EVRISSTKMIEREINSVNDPILIVSRKALHSGNFQSTPIGVS 422
QY 405 DNTLRLAIAIKLMAPOSELVNDPYNNGILPSNLSSGRNPFLDGRKSGELIAMAASYSEL 464
DB 423 EKTRLGLAOIGKLNFTQLEMLNAGMNRGLPSCLA--AEDESLSHCKGDLIAAAAYTSEL 481
QY 465 OFLANPVNHHQSAOHQOVNSLGLISARKTAAVLDLILMSTYVVALCOSIDLRIHLE 524
DB 482 GHLANPVTHVQAPMAQAVNSLALLISARKTESNVLSLLATHLHYCVLQALIDIRAXE 541
QY 525 ENMKSTVKNVTSQVAKKVLTMGVNCELHPSRFCEKDLIRVVDREYVAYIDVCSGYPL 584
DB 542 FEFKK-----QFGPA-----IVSLIDQH-----GSAMT 565
QY 585 MOKLRVLDHALNNGETEKNTNIFQKIATFEELKVLILPKREYGV-----632
DB 566 GSNLDELIV--KVNKTIAKRLKQTNSTYDLVRMHDAFSFAGTYVEVLSTSLAAVNA 624
QY 633 -RIAYENDTLISPNRIK-----ACRSYP-----LYRVRRELG---RG--FL 668
DB 625 WKVAAEASALISLTROVRETFMSASTSPALSYLSPRTQILVAVREBELGVAKRGDVL 684
QY 669 TGEKVTSPGEEDRVFTAMCKGQIIDLLECL 700

Db 685 GQEVTT-GSNVSKTYFAIKSGRIINNVLKML 715

RESULT 15

US-09-765-873A-32
 ? Sequence 32, Application US/09765873A
 ? Patent No. US20010053847A1
 ? GENERAL INFORMATION:
 ? APPLICANT: Tang, Xiao-Song
 ? TITLE OF INVENTION: BIOPROTECTION OF PAPA-HITP-XY-INNAMIN Acid
 ? FILE REFERENCE: H01009 US CIP
 ? CURRENT APPLICATION NUMBER: US/09/765,873A
 ? PRIORITY FILING DATE: 2001-01-19
 ? PRIOR APPLICATION NUMBER: US 09/627,216
 ? PRIOR FILING DATE: 2000-07-27
 ? PRIOR APPLICATION NUMBER: US 60/147,719
 ? NUMBER OF SEQ ID NOS: 38
 ? SOFTWARE: Microsoft Office 97
 ? SEQ ID NO 32
 ? LENGTH: 686
 ? TYPE: PRT
 ? ORGANISM: mutant from Rhodotorula glutinis
 ? US-09-765-873A-32

Query Match 27.1% Score 988; DB 10; Length 686;

Best Local Similarity 36.2%; Pred No. 9, 6e-68;
 Matches 265; Conservative 121; Mismatches 240; Indels 106; Gaps 22.

QY 28 ALTGSHL-----DEYKKVAFPRKIVVKLGGFTLLVSVAQIAAANDPIYKVLSE 79
 Db 1 AVAGSHLPTQYQYDIYERKMLAAPDSTLELDGYSINLGDV--VSAARKGRPYRVKSD 58
 QY 80 ARAGVYKASSDMYMSNMKKTDSYGVTTGPGATSHPTKQSGALQKELIRFLNAGT----- 135
 Db 59 EIRSKIDKSVFPIKRSQISMSV--YGVTTGPGGADPTREDAISLOKALIRHQLCGVLPSS 116
 QY 136 -----FGKGTETSTLPHSATPAAMIVRINTLGGYSQIFELIETKFLNNITPCLP 190
 Db 117 FDSFRLGRLNS--LPLEVYKGMATIRVNSLTGRHSVRLVLEALTNFLNIGITPIY 174
 QY 191 LMGTTTASGDLVPISYIAGLLTGRPNKA--VGPIC--FVINAERKAFAGVGGCFRLLQ 247
 Db 175 LRGTTASGDSPLSYIAAIGSHPSKVIHVHEGKEKILYAREMALFNLDP--VVLGP 232
 QY 248 KEGLLVNGTAVSGNAGMVLFDANVLALLSEVLATFAEYMOGR--EFTDHLTKLKH 306
 Db 233 KRGDGLVNGTAVSASMATLALHDAHMLSLISOSTIATMVEAMVGHAGSPHPIHDTYRPH 292
 QY 307 PQGLEMAAMEYILDGSDY--VKAAGKYHEMDPLVKPKDARYALKTSFQWLGPOI-EYIRS 364
 Db 293 PTOIEVANIRKLEGSREFAVHHEEYKVKDDEGILRDYRPLRTSPQWLGPLVSDLIHA 352
 QY 365 STKMTERKINSYNDPILIDVSKNKALHOGNOCGTPICVSMNTRKLAIAAIGKIMFAQFS 424
 Db 353 HAVLTLEGQSTDPPLIDVENKTSIHGNGFOAAAVANTMEKTRLGIAIGKLNFTOLTE 412
 QY 425 LVNDEYNNGLSPNSLGSFNPSTLYGFKGSEYFAMASYCELDQFLANPVTNNIVUSAEQHNOD 484
 Db 413 MNAQMNNGIIPSCIA-ARDPISLYHCKGLIDIAAAVYTHOHLANIPVTHVQPAEMANQA 471
 QY 485 VNSLGLISARKTAENVILKLTMSSTYVALGQSIDLRLHLEENKSTVKNTVSQYAKKVL 544
 Db 472 VNSLLISARRTTESNDVLSLLATHLYCVLAIDLRAIEFEFK----- 516
 QY 545 MGVNCHLSHSRCEKDLIRVDREYFAYIDDVSGTYPIIMOKIROYVJDAHINNGETEK 604
 Db 517 -----QFGPA-----IVSLIDHF-----GSAMTGSNLRDELVE-KVAKTLAKR 554
 QY 605 NNTSIFOKIAIEEELKVLPRKEGV-----RIAYENDTSLIPNRK--- 648
 Db 555 LKQTSYDILVPRWHDAFSAAGTAVFVLIJSTSLIAAVNAWKVAASASISITPROVRETF 614

QY 649 ---ACKSY-----LYREFREHIG---KG--PIIGCKVTSPOGHPURKVTAMC 688
 Db 615 WSAASTSSPALSYLSPPTQTLYAFVPEHIAVKAPRGNVPIKQEVTT-GSNVSKTYEALR 673
 QY 689 KQGLIDPIPLPGL 700
 Db 674 SGRIINNVLKML 685

Search completed: March 29, 2003, 02:54:43
 Job time: 184 secs

GenCore version 5.1.4 P5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 29, 2003, 02:05:46 : Search time 49 seconds
(without alignments)
1394.930 Million cell updates/sec

Title: US-09-964-992A-1
Perfect score: 3644
Sequence: 1 MEN:NIHVN:VYNEI/TKDPL ITDPLECT/GCWNGEPLPIC 711

Scoring table: HIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73: *
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	DB ID	Description
1	3206	88.0	667	2	T12749
2	3186.5	87.4	716	2	S04463
3	3186.5	87.4	718	2	S48726
4	3185.5	87.4	716	2	S48725
5	3171.5	87.0	708	2	T14295
6	3149	86.4	715	2	S66343
7	3118.5	85.6	712	2	T03663
8	3116	85.5	715	2	J02265
9	3109.5	85.3	712	2	T01858
10	3077.5	84.5	710	2	J05872
11	3070	84.2	723	2	S25901
12	3068	84.2	713	2	S24991
13	3067.5	84.2	721	2	A44133
14	3067.5	84.2	725	2	S17444
15	3058	83.9	712	2	J04127
16	3051	83.7	705	2	J05873
17	3049	83.7	720	2	S21174
18	3045.5	83.6	725	2	G84787
19	3039.5	83.4	725	2	S52990
20	3004	82.4	717	2	T146172
21	2988	82.0	708	2	T109399
22	2985.5	81.9	710	2	S60042
23	2953.5	81.1	707	2	S29629
24	2730.5	74.9	700	2	T06545
25	2671	73.3	619	2	S52632
26	2632.5	72.2	710	2	S66313
27	2598	71.3	582	2	S60043
28	2590	71.1	710	2	S04128
29	2578.5	70.8	695	2	S52992

ALIGNMENTS

30	2555.5	70.1	716	2	S28185	phenylalanine amno
31	2538	69.6	590	2	S70916	phenylalanine amno
32	2496	68.5	701	2	S06475	phenylalanine amno
33	2373.5	65.1	754	2	T09777	phenylalanine amno
34	2220	60.9	505	2	A24727	phenylalanine amno
35	2220	60.9	506	2	S04129	phenylalanine amno
36	2184	59.4	549	2	T05970	phenylalanine amno
37	1948	53.5	475	2	T05966	phenylalanine amno
38	1887	51.8	497	2	T05968	probable phenyla
39	1867	51.2	416	2	J01070	phenylalanine amno
40	1368	37.5	339	2	T06546	phenylalanine amno
41	994.5	27.3	716	2	A56628	phenylalanine amno
42	979	26.9	235	2	S25538	phenylalanine amno
43	977	26.8	693	2	A29607	phenylalanine amno
44	971.5	26.7	713	2	S01999	phenylalanine amno
45	540.5	14.8	511	2	E82228	histidine ammonia

RESULT 1

T12749

phenylalanine ammonia-lyase (EC 4.3.1.5) - common sunflower

C:Species: Helianthus annuus (common sunflower)

C:Date: 13 Aug 1999 #sequence_revision 13 Aug 1999 #text_change 29 Jun 2000

C:Accession: T12749

R:Maxeyrat, F.A.

A:Submitted to the EMBL Data Library, February 1998

A:Reference number: Z17579

A:Accession: T12749

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-667 <MAX>

A:Cross-references: EMBL: Y12461

C:Genetics:

A:Gene: PAL

C:Superfamily: histidine ammonia-lyase

C:Keywords: ammonia-lyase; carbon-nitrogen lyase

F:197-199/Cross-link: 5-imidazolinone (Ala-Gly) #status predicted

F:198/Modified site: dehydroalanine (Ser) #status predicted

Query Match

Best Local Similarity 88.0%: Score 3206; DB 2; Length 667;

Matches 667, Conservative 13; Mismatches 24; Indels 0; Gaps 0;

QY	1	MENGNHVNCGVNEHLCIKDPLNNGVAFAHATGSHIDVKKWAEKKVVKIGCHETIVSO	60
DB	1	MENGNHVNCGVNEHLCIKDPLNNGVAFAHATGSHIDVKKWAEKKVVKIGCHETIVSO	60
QY	61	VATIAAANISIVVYHLSAAKATVKAASAWMEAMNGGHSYVYTGFAVSHPIKNG	120
DB	61	VATISAGGQNNVYVHLSAAKATVKAASAWMEAMNGGHSYVYTGFAVSHPIKNG	120
QY	121	GALQKELREFLVNAGIFGNGTETSHTLPHSATRAAMIVHINTLQVYSAINLELEATIKF	180
DB	121	GALQKELREFLVNAGIFGNGTETSHTLPHSATRAAMIVHINTLQVYSAINLELEATIKF	180
QY	181	INNNITPCGIPGICITIASGLWPSYVIAALIGRNSAAVGTGVINAAKAPAAATVEG	240
DB	181	INNNITPCGIPGIPGICITIASGLWPSYVIAALIGRNSAAVGTGVINAAKAPAAATVEG	240
QY	241	GPEHQPQKGLAIVNCTAAGSCMASWVLPDANVLAISSEVSAIPAEVMQCKPPEPTHTT	300
DB	241	GPEHQPQKGLAIVNCTAAGSCMASWVLPDANVLAISSEVSAIPAEVMQCKPPEPTHTT	300
QY	301	HKRKHHPGJLEAAALMEYTLPSDVKAAQVHEMLPKVKKKQVYALRTSPQDLSPQIE	360
DB	301	HKRKHHPGJLEAAALMEYTLPSDVKAAQVHEMLPKVKKKQVYALRTSPQDLSPQIE	360
QY	361	VIRSTKMKIEKEINSDVNPPLIVVSHNKAHGGCNQGITGVSMNINHLAIAALGKIMPA	420
DB	361	VIRSAIKMKIEKEINSDVNPPLIVVSHNKAHGGCNQGITGVSMNINHLAIAALGKIMPA	420

QY 421 QPSELVNDYNNGLPSNLSCGRNPSLDYGFCKGEIAMASYCSHLOFLANPTVTHNVSARQ 480
 Db 421 QPSELVNDYNNGLPSNLSCGRNPSLDYGFCKGEIAMASYCSHLOFLANPTVTHNVSARQ 480
 QY 481 HNODVNSLGLISARKTAFAVDILKIMSSYLVALCOSIDIRHLLENKSTVKNVSOVAK 540
 Db 481 HNODVNSLGLISARKTAFAVDILKIMSSYLVALCOSIDIRHLLENKSTVKNVSOVAK 540
 QY 541 KYLTGNGVCELHPSPFCEKDLRYVDREYFAYTDDVCSGTTPLMOKLROYLVDAHALNG 600
 Db 541 KYLTGNGVCELHPSPFCEKDLRYVDREYFAYTDDVCSGTTPLMOKLROYLVDAHALNG 600
 QY 601 ETEKNTNSIFOKIATFEDELKALLPKEVESVRAVFNENGMISIPNRKACRSYPLRYFV 660
 Db 601 ETEKNTNSIFOKIATFEDELKALLPKEVESVRAVFNENGMISIPNRKACRSYPLRYFV 660
 QY 661 EELG 664
 Db 661 EELG 664

RESULT 2

S04463
 Phenylalanine ammonia-lyase (EC 4.3.1.5) - parsley
 C:Species: Petroselinum crispum (parsley)
 C:Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 21-May-1999
 C:Accession: S04463
 R:Lois, R.; Dietrich, A.; Hahlbrock, K.; Schulz, W.
 EMBL J. 8, 1641-1648, 1989
 A:Title: A phenylalanine ammonia-lyase gene from parsley: structure, regulation and identification
 A:Reference number: S04463, MUID:89356635, PMID:2767049
 A:Accession: S04463
 A:Molecule type: DNA
 A:Residues: 1 716 <LOI>
 A:Cross-references: EMBL:X15473
 C:Genetics:
 A:Introns: 136/2
 C:Superfamily: histidine ammonia-lyase
 C:Keywords: ammonia-lyase; carbon-nitrogen lyase
 F:202/204/Cross-link: 5-imidazolone (Ala-Gly) #status predicted
 F:203/Modified site: dehydroalanine (Ser) #status predicted

Query Match 87.4%; Score 3186.5; DB 2; Length 716;
 Best Local Similarity 86.2%; Pred. No. 2.3e-196;
 Matches 620; Conservative 46; Mismatches 42; Indels 11; Gaps 4;
 QY 1 MENCN-----HVGCVNVLCTK--DPLNMVAAEALTGSHLDEYKKMAEPRKPVYKLS 52
 Db 1 MENCN-----HVGCVNVLCTK--DPLNMVAAEALTGSHLDEYKKMAEPRKPVYKLS 52
 QY 53 GETLVSQVAGITAAANDSDYKVELSEARAGVAKSSDWYVESNKKGTDSYGVTTGPGAT 112
 Db 53 GETLVSQVAGITAAANDSDYKVELSEARAGVAKSSDWYVESNKKGTDSYGVTTGPGAT 112
 QY 61 GETLISQVAAI-SARDSCGVTVLSEARAGVAKSSDWYVESNKKGTDSYGVTTGPGAT 119
 Db 61 GETLISQVAAI-SARDSCGVTVLSEARAGVAKSSDWYVESNKKGTDSYGVTTGPGAT 119
 QY 113 SHRRITKOGALOKELIRPLNAGIFGNCETSHLPHSATRAAMIVRINTLLQYSGIRRE 172
 Db 113 SHRRITKOGALOKELIRPLNAGIFGNCETSHLPHSATRAAMIVRINTLLQYSGIRRE 172
 QY 120 SHRRITKOGALOKELIRPLNAGIFGNSD--NTLPHSATRAAMIVRINTLLQYSGIRRE 177
 Db 120 SHRRITKOGALOKELIRPLNAGIFGNSD--NTLPHSATRAAMIVRINTLLQYSGIRRE 177
 QY 173 ILLEATKFLNNITPPCLPRTITASGDLVPLSTIAGLLTGTRNSKAVGPTGVLNAFKA 232
 Db 173 ILLEATKFLNNITPPCLPRTITASGDLVPLSTIAGLLTGTRNSKAVGPTGVLNAFKA 232
 QY 178 ILLEATKFLNNITPPCLPRTITASGDLVPLSTIAGLLTGTRNSKAVGPTGVLNAFKA 237
 Db 178 ILLEATKFLNNITPPCLPRTITASGDLVPLSTIAGLLTGTRNSKAVGPTGVLNAFKA 237
 QY 233 FAAGVGEGETFELQPKELALVNGTAVSGMASVLFDAVNLALLSEVLSATFAEYMOG 292
 Db 233 FAAGVGEGETFELQPKELALVNGTAVSGMASVLFDAVNLALLSEVLSATFAEYMOG 292
 QY 238 FKLAGVGEGETFELQPKELALVNGTAVSGMASVLFDAVNLALLSEVLSATFAEYMOG 297
 Db 238 FKLAGVGEGETFELQPKELALVNGTAVSGMASVLFDAVNLALLSEVLSATFAEYMOG 297
 QY 293 PEFTDHLTHKLKHPGQIEAAIMEYILDSGDYKAAQKQVHNDPLOKPRQDYALRTSP 352
 Db 293 PEFTDHLTHKLKHPGQIEAAIMEYILDSGDYKAAQKQVHNDPLOKPRQDYALRTSP 352
 QY 298 PEFTDHLTHKLKHPGQIEAAIMEYILDSGDYKAAQKQVHNDPLOKPRQDYALRTSP 357
 Db 298 PEFTDHLTHKLKHPGQIEAAIMEYILDSGDYKAAQKQVHNDPLOKPRQDYALRTSP 357
 QY 353 QMLGQIEVILKSSITKMLEREINSVNDPLIDVSNKALHCGNFQGTPIGVSMNTRIALA 412
 Db 353 QMLGQIEVILKSSITKMLEREINSVNDPLIDVSNKALHCGNFQGTPIGVSMNTRIALA 412

Db 358 QMLGQIEVILKSSITKMLEREINSVNDPLIDVSNKALHCGNFQGTPIGVSMNTRIALA 417
 QY 413 AIGKLMFAQSESELVNDYNNGLPSNLSCGRNPSLDYGFCKGEIAMASYCSHLOFLANPT 472
 Db 413 AIGKLMFAQSESELVNDYNNGLPSNLSCGRNPSLDYGFCKGEIAMASYCSHLOFLANPT 472
 QY 473 HNODVNSLGLISARKTAFAVDILKIMSSYLVALCOSIDIRHLLENKSTVKNVSOVAK 532
 Db 473 HNODVNSLGLISARKTAFAVDILKIMSSYLVALCOSIDIRHLLENKSTVKNVSOVAK 532
 QY 478 HNODVNSLGLISARKTAFAVDILKIMSSYLVALCOSIDIRHLLENKSTVKNVSOVAK 537
 Db 478 HNODVNSLGLISARKTAFAVDILKIMSSYLVALCOSIDIRHLLENKSTVKNVSOVAK 537
 QY 533 NTVSOVAKKVLTVGNGVCELHPSPFCEKDLRYVDREYFAYTDDVCSGTTPLMOKLROYL 592
 Db 533 NTVSOVAKKVLTVGNGVCELHPSPFCEKDLRYVDREYFAYTDDVCSGTTPLMOKLROYL 592
 QY 593 VDHALLNCGETKNTNSIFOKIATFEDELKALLPKEVESVRAVFNENGMISIPNRKACRSYPLRYFV 652
 Db 593 VDHALLNCGETKNTNSIFOKIATFEDELKALLPKEVESVRAVFNENGMISIPNRKACRSYPLRYFV 652
 QY 653 YPLKFAVAKKVLTVGNGVCELHPSPFCEKDLRYVDREYFAYTDDVCSGTTPLMOKLROYL 711
 Db 653 YPLKFAVAKKVLTVGNGVCELHPSPFCEKDLRYVDREYFAYTDDVCSGTTPLMOKLROYL 711

RESULT 3

S48726
 Phenylalanine ammonia-lyase (EC 4.3.1.5) 3 - parsley
 C:Species: Petroselinum crispum (parsley)
 C:Date: 18-Oct-1995 #sequence_revision 12-Apr-1996 #text_change 16-Jul-1999
 C:Accession: S48726; S56035
 R:Appert, C.; Logemann, E.; Hahlbrock, K.; Schmid, J.; Amrhein, N.
 Eur. J. Biochem. 225, 491-499, 1994
 A:Title: Structural and catalytic properties of the four phenylalanine ammonia-lyase
 A:Reference number: S48725; MUID:95010141; PMID:7925471
 A:Accession: S48726
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-718 <APP1>
 A:Cross-references: EMBL:X81159
 R:Appert, C.
 Submitted to the EMBL Data Library, August 1994
 A:Reference number: S56035
 A:Accession: S56035
 A:Molecule type: mRNA
 A:Residues: 1-58, '6'-718 <APP2>
 A:Cross-references: EMBL:X81159; NID:9535007, PIDN:CAA57057.1; FID:9535008
 C:Genetics:
 A:Gene: PAL3
 C:Superfamily: histidine ammonia-lyase
 C:Keywords: ammonia-lyase; carbon-nitrogen lyase
 F:204/206/Cross-link: 5-imidazolone (Ala-Gly) #status predicted
 F:205/Modified site: dehydroalanine (Ser) #status predicted

Query Match 87.4%; Score 3186.5; DB 2; Length 718;
 Best Local Similarity 87.2%; Pred. No. 2.3e-196;
 Matches 621; Conservative 46; Mismatches 40; Indels 5; Gaps 4;
 QY 3 NGNHNHGVNDELCTK--DPLNMVAAEALTGSHLDEYKKMAEPRKPVYKLS 60
 Db 3 NGNHNHGVNDELCTK--DPLNMVAAEALTGSHLDEYKKMAEPRKPVYKLS 60
 QY 9 NG-HANGSGDLCKMKEDPLNMVAAEALTGSHLDEYKKMAEPRKPVYKLS 67
 Db 9 NG-HANGSGDLCKMKEDPLNMVAAEALTGSHLDEYKKMAEPRKPVYKLS 67
 QY 61 VAGIAAANDSDYKVELSEARAGVAKSSDWYVESNKKGTDSYGVTTGPGATSHRPTKOG 120
 Db 61 VAGIAAANDSDYKVELSEARAGVAKSSDWYVESNKKGTDSYGVTTGPGATSHRPTKOG 120
 QY 68 VAAISARDDSG-VVEVLESEARAGVAKSSDWYVESNKKGTDSYGVTTGPGATSHRPTKOG 126
 Db 68 VAAISARDDSG-VVEVLESEARAGVAKSSDWYVESNKKGTDSYGVTTGPGATSHRPTKOG 126
 QY 121 GALKKELIRFLNAGIFGNSD--NTLPHSATRAAMIVRINTLLQYSGIRRELEAIFK 179
 Db 121 GALKKELIRFLNAGIFGNSD--NTLPHSATRAAMIVRINTLLQYSGIRRELEAIFK 179
 QY 127 GALKKELIRFLNAGIFGNSD--NTLPHSATRAAMIVRINTLLQYSGIRRELEAIFK 186
 Db 127 GALKKELIRFLNAGIFGNSD--NTLPHSATRAAMIVRINTLLQYSGIRRELEAIFK 186
 QY 180 FLNNITPPCLPRTITASGDLVPLSTIAGLLTGTRNSKAVGPTGVLNAFKAFAAGVE 239
 Db 180 FLNNITPPCLPRTITASGDLVPLSTIAGLLTGTRNSKAVGPTGVLNAFKAFAAGVE 239
 QY 187 FLNNITPPCLPRTITASGDLVPLSTIAGLLTGTRNSKAVGPTGVLNAFKAFAAGVE 246
 Db 187 FLNNITPPCLPRTITASGDLVPLSTIAGLLTGTRNSKAVGPTGVLNAFKAFAAGVE 246
 QY 240 GCFEELQPKELALVNGTAVSGMASVLFDAVNLALLSEVLSATFAEYMOGKPEFTDHL 299
 Db 240 GCFEELQPKELALVNGTAVSGMASVLFDAVNLALLSEVLSATFAEYMOGKPEFTDHL 299


```

|||||
Db 247 GCFEELQPKKGLALVNGTAVGSGMASMVLFRANILLAVAFVMSALFAVEMQCKPFFTHI 306
Qy 300 THIKIKHHPQVLEAAIMEYILDSSIVYKAAVKHEMLPULQKRCQCPVALFTSPYMLSPD1 354
Db 307 THIKIKHHPQVLEAAIMEYILDSSIVYKAAVKHEMLPULQKRCQCPVALFTSPYMLSPD1 366
Qy 360 EVLRSSTKMERELENSVNDPLDVSFNKATIMFGNFOSSPTGVSMONIRLALIAIGKLMF 419
Db 367 EVLRSSTKMERELENSVNDPLDVSFNKATIMFGNFOSSPTGVSMONIRLALIAIGKLMF 426
Qy 420 AQSSELVNDPYNNGDPSNLSGGRNPSIDYGFQKGEIAMAASVSCSELQIAPNPNHVSAR 479
Db 427 AQSSELVNDPYNNGDPSNLSGGRNPSIDYGFQKGEIAMAASVSCSELQIAPNPNHVSAR 486
Qy 480 QINDVNSLILSAKTAFAENVILKLMSSITYVALQVSTDLKHLEBNKSYVKNTVSYVA 539
Db 487 QINDVNSLILSAKTAFAENVILKLMSSITYVALQVSTDLKHLEBNKSYVKNTVSYVA 546
Qy 540 KKVITMGVNGELHPSRPECEDLLRVDRKYVFAVIDVCSGYPIPMOKLQVLYDHALNN 599
Db 547 KKVITMGVNGELHPSRPECEDLLRVDRKYVFAVIDVCSGYPIPMOKLQVLYDHALNN 606
Qy 600 GETTKNTNISIFOKIAFFHEELKVLIPKKEGVRIAYINDTILSPNRIKACRSPLRYRV 659
Db 607 GDKRKNLSTISFOKIAAFDEELKALLPKKEVETARAALISGNHAIIPNRIKCKRSPLYKRV 666
Qy 660 REELGRGFTGKEVTSFGEEFDPVFTAMKQGITDPLELCTGFWNCEPLTIC 711
Db 667 REELGRGFTGKEVTSFGEEFDPVFTAMKQGITDPLELCTGFWNCEPLTIC 718

```

RESULT 4

S48725
phenylalanine ammonia-lyase (EC 4.3.1.5) 2 - parsley
C:Species: Petroselinum crispum (parsley)
C>Date: 28 Oct 1995 #sequence_revision 10 Nov 1995 #excl_change 1b-Jul-1999
C:Accession: S48725
R:Appert, C.; Lodemann, E.; Hahlbrock, K.; Schmid, J.; Amphelin, N
Eur. J. Biochem. 225, 491-499, 1994
A:Title: Structural and catalytic properties of the four phenylalanine ammonia-lyase iso
A:Reference number: S48725, M01D.95010141, PMID.7925471
A:Accession: S48725
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1,716 - APP
A:Cross-references: EMBL:X81158; NID:q514892; PIRN:CMA57056 1; PTD:q534893
C:Genetics:
A:Gene: PAL2
C:Superfamily: histidine ammonia-lyase
C:Keywords: ammonia-lyase; carbon-nitrogen lyase
F:102-204/cross-link: 5 imidazolinone (Ala-Gly) #status predicted
F:203/Modified site: dehydroalanine (Ser) #status predicted

Query Match 87.4%, Score 3185.5, DB 2; Length 716;
Best Local Similarity: 86.2%, Pred. No. 26e-196;
Matches 620; Conservative 45; Mismatches 43; Indels 11; Gaps 4;

```

Qy 1 MONGNCAITNGVNGELCIK--DPLNMGVAAEALTGSILDEYKKVAERKRPVKLG 52
Db 1 MONGNCAITNGVNGELCIK--DPLNMGVAAEALTGSILDEYKKVAERKRPVKLG 60
Qy 53 GFTLTQVAVGATAANDSDTVKVELSEARAVKASSDWMESNNKGTDSYVTTGFGAT 112
Db 61 GFTLTQVAVGATAANDSDTVKVELSEARAVKASSDWMESNNKGTDSYVTTGFGAT 119
Qy 113 SHRTKGGALQKELIKFLNAGIFGNGTFTSHIIPHSATRAAMIVRIINLQGYSGIHP 172
Db 120 SHRTKGGALQKELIKFLNAGIFGNGTFTSHIIPHSATRAAMIVRIINLQGYSGIHP 177
Qy 173 ILEAITFELNNNTTICPLPLRGITTSAGDLVPLSYIAGLLTGPPNKRKAVGPTGEVLA 232
Db 178 ILEAITFELNNNTTICPLPLRGITTSAGDLVPLSYIAGLLTGPPNKRKAVGPTGEVLA 237

```

```

Qy 233 FAAGVNDSTFELQKRGALVNGTAVGSGMASMVLFRANVILLSEVLSALFAVEMQCK 292
Db 238 FKLAVNDSTFELQKRGALVNGTAVGSGMASMVLFRANVILLSEVLSALFAVEMQCK 297
Qy 293 PEPTDHTLTHIKKHHPQVLEAAIMEYILDSSIVYKAAVKHEMLPULQKRCQCPVALFTSP 352
Db 298 PEPTDHTLTHIKKHHPQVLEAAIMEYILDSSIVYKAAVKHEMLPULQKRCQCPVALFTSP 357
Qy 353 QWLQPOLVITPSSTKMTIEPEINSVNDPLDVSFNKATIMFGNFOSSPTGVSMONIRLALIA 412
Db 358 QWLQPOLVITPSSTKMTIEPEINSVNDPLDVSFNKATIMFGNFOSSPTGVSMONIRLALIA 417
Qy 413 ALCKIMAPQSEIIVNDPYNNGDPSNLSGGRNPSIDYGFQKGEIAMAASVSCSELQIAPNPN 472
Db 418 ALCKIMAPQSEIIVNDPYNNGDPSNLSGGRNPSIDYGFQKGEIAMAASVSCSELQIAPNPN 477
Qy 473 NHVSARQNDVNSLILSAKTAFAENVILKLMSSITYVALQVSTDLKHLEBNKSYVKNTVSYVA 532
Db 478 NHVSARQNDVNSLILSAKTAFAENVILKLMSSITYVALQVSTDLKHLEBNKSYVKNTVSYVA 537
Qy 533 NTVSQAQKVLTVNGVNGELHPSRPECEDLLRVDRKYVFAVIDVCSGYPIPMOKLQVLYD 592
Db 538 NTVSQAQKVLTVNGVNGELHPSRPECEDLLRVDRKYVFAVIDVCSGYPIPMOKLQVLYD 597
Qy 593 VDHALNNGFTGKKNNTISIFOKIAFFHEELKVLIPKKEGVRIAYINDTILSPNRIKACRS 652
Db 598 VDHALNNGFTGKKNNTISIFOKIAFFHEELKVLIPKKEGVRIAYINDTILSPNRIKACRS 657
Qy 653 YPLRYVREELQKGLITGKRVKTSFGEEFDPVFTAMKQGITDPLELCTGFWNCEPLTIC 711
Db 658 YPLRYVREELQKGLITGKRVKTSFGEEFDPVFTAMKQGITDPLELCTGFWNCEPLTIC 716

```

RESULT 5

T14295
phenylalanine ammonia-lyase (EC 4.3.1.5) - carrot
C:Species: Daucus carota (carrot)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #excl_change 21-Jul-2000
C:Accession: T14295
R:Takeuchi, Y.; Ozeki, Y.; Yoshida, K.
Photochem. Photobiol. 66, 464-470, 1997
A:Title: Action spectrum for induction of promoter activity of phenylalanine ammonia
A:Reference number: 217966, M01D.97478956, PMID.9337617
A:Accession: T14295
A:Status: Preliminary, translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1,708 - TAK
A:Cross-references: EMBL:U85850; NID:q2618589; PIRN:BA23367.1; PTD:q2618590
A:Experimental source: strain: Kuradagousun
C:Genetics:
A:Gene: PAL1
A:Introns: 126/2
C:Function:
A:Description: catalyzes the deamination of L-phenylalanine to trans-cinnamic acid
A:Pathway: phenylpropanoid metabolism
C:Superfamily: histidine ammonia-lyase
C:Keywords: ammonia-lyase; carbon-nitrogen lyase
F:104-196/cross-link: 5 imidazolinone (Ala-Gly) #status predicted
F:195/Modified site: dehydroalanine (Ser) #status predicted

Query Match 87.0%, Score 3171.5, DB 2; Length 708;
Best Local Similarity: 87.2%, Pred. No. 2e-195;
Matches 619; Conservative 41; Mismatches 43; Indels 7; Gaps 3;

```

Qy 2 ENCHRVNGVNGELCI-KDPLNMGVAAEALTGSILDEYKKVAERKRPVKLGCELTIVSO 60
Db 4 ENKRVNIG--NGICMOKDPLNMGMAAKALTGSILDEYKKVAERKRPVKLGCELTIVSO 61
Qy 61 VAGTAANDSDTVKVELSEARAVKASSDWMESNNKGTDSYVTTGFGATSHRTKGG 120
Db 62 VAAIAAG----SVKVLASARAVKASSDWMESNNKGTDSYVTTGFGATSHRTKGG 117

```

```

QY 121 GALKELIRFLNAGIFNGTETSHLPASATRAAMIVKINTLLQSGISGIRETEATEATKE 180
Db 118 GAIQKELIRFLNAGIFNGTETSHLPASATRAAMIVKINTLLQSGISGIRETEATEATKE 177
QY 181 LNNITPCLPLRGITITASGDIPLVLSYIAGLLTGRPNKAVGPTGEVIMAKKAPAAAGVGC 240
Db 178 LNNITPCLPLRGITITASGDIPLVLSYIAGLLTGRPNKAVGPTGEVIMAKKAPAAAGVGC 237
QY 241 GFEELQKREGIALVNGTAVGSMASVYLFEDANVALLSEVLSAIFPAVMOGKREPTDHLT 300
Db 238 GFEELQKREGIALVNGTAVGSMASVYLFEDANVALLSEVLSAIFPAVMOGKREPTDHLT 297
QY 301 HKLHHHPQIEAAAMEYILDGSDYKAAQKHEMDPLQKPKODRYALRTSPQWLGPOIE 360
Db 298 HKLHHHPQIEAAAMEYILDGSDYKAAQKHEMDPLQKPKODRYALRTSPQWLGPOIE 357
QY 361 VIRSTKMIEREINSVNDNPLIDVSRNKKALHGNFQCTPIGVSMQNTIRIALAAICGLMA 420
Db 358 VIRSTKMIEREINSVNDNPLIDVSRNKKALHGNFQCTPIGVSMQNTIRIALAAICGLMA 417
QY 421 QFSELVNDVFNNGLPSPNLSGGRNPSLDYGFKGEIIMASYSCELOFLANPVTNHVQSAQ 480
Db 418 QFSELVNDVFNNGLPSPNLSGGRNPSLDYGFKGEIIMASYSCELOFLANPVTNHVQSAQ 477
QY 481 HNUVNSLGLISARKTAEAVDILKLMSSYVALCOSIDLRLLENNKSTVKNVSQVAK 540
Db 478 HNUVNSLGLISARKTAEAVDILKLMSSYVALCOSIDLRLLENNKSTVKNVSQVAK 537
QY 541 KVLIMGVNGLHPSPRCEKDLRYVDREYFAYIDVCGSTYPLMKIHOVLDHALNNG 600
Db 538 KVLIMGVNGLHPSPRCEKDLRYVDREYFAYIDVCGSTYPLMKIHOVLDHALNNG 597
QY 601 ETEKNTNTSIFOKIATFEELKVLIPKEVEGVRIAYENDTJSPNRIKACRSYPLRFAV 660
Db 598 ETEKNTNTSIFOKIATFEELKVLIPKEVEGVRIAYENDTJSPNRIKACRSYPLRFAV 657
QY 661 EELRGFTLGKVTSPGEEFDRTFTAMCKGQIIDLPLECLGGNNGPLPI 710
Db 658 EELRGFTLGKVTSPGEEFDRTFTAMCKGQIIDLPLECLGGNNGPLPI 707

RESULT 6
S66343
phenylalanine ammonia-lyase (EC 4.3.1.5) 1 - common tobacco
C:Species: Nicotiana tabacum (common tobacco)
C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 16-Jul-1999
C:Accession: S66343
R:Fukasawa-Akada, T.; Kung, S.; Watson, J.C.
Plant Mol Biol 30, 711-722, 1996
A:Title: Phenylalanine ammonia-lyase gene structure, expression, and evolution in Nicotiana
A:Reference number: S66343; MUID:96194455; PMID:86244404
A:Accession: S66343
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-715 <EFUK>
A:Cross-references: EMBL:M84466; NID:q170349; PION:AAA34122.1; PTD:q170350
C:Genetics:
A:Gene: PAL1
A:Introns: 13/2
A:Superfamily: histidine ammonia-lyase
C:Keywords: ammonia-lyase; carbon-nitrogen lyase
F:201-203/Cross-link: 5-imidazolone (Ala-Gly) #status predicted
F:202/Modified site: dehydroalanine (Ser) #status predicted

Query Match 86.4%; Score 3149; DB 2; Length 715;
Best Local Similarity 85.3%; Pred. No. 5,7e-194;
Matches 610; Conservative 45; Mismatches 56; Indels 4; Gaps 3;

QY 1 MEMNHVNGVNV-ELCTK--DPLNMGVAALNLSHLDVKKKVAARFRKRPVYKLGCTTR 57
Db 1 MASGHNVGGENFELCKKSADPLNMEAAASLGSLSHLDVKKKVAARFRKRPVYKLGCTTR 60

QY 58 VSQVAGIAAANDS-DIVKVELSEARAGVAKSSDWVESMKNKGTSDYGVTTGGATSHRR 116

```

```

Db 61 VAQVAAIAAVNDKSNKGVKVELSEARAGVAKSSDWVMDMKNKGTSDYGVTTGGATSHRR 120
QY 117 TKGGALQKELIRFLNAGIFNGTETSHLPASATRAAMIVKINTLLQSGISGIRETEATEATKE 176
Db 121 TKGGALQKELIRFLNAGIFNGTETSHLPASATRAAMIVKINTLLQSGISGIRETEATEATKE 180
QY 177 ITRFLNNITPCLPLRGITITASGDIPLVLSYIAGLLTGRPNKAVGPTGEVIMAKKAPAAAGVGC 236
Db 181 ITRFLNNITPCLPLRGITITASGDIPLVLSYIAGLLTGRPNKAVGPTGEVIMAKKAPAAAGVGC 240
QY 237 GVEGFEELQKREGIALVNGTAVGSMASVYLFEDANVALLSEVLSAIFPAVMOGKREPTDHLT 296
Db 241 GVEGFEELQKREGIALVNGTAVGSMASVYLFEDANVALLSEVLSAIFPAVMOGKREPTDHLT 300
QY 297 DHLTTHKLHHHPQIEAAAMEYILDGSDYKAAQKHEMDPLQKPKODRYALRTSPQWLGPOIE 356
Db 301 DHLTTHKLHHHPQIEAAAMEYILDGSDYKAAQKHEMDPLQKPKODRYALRTSPQWLGPOIE 360
QY 357 POIEVIRSTKMIEREINSVNDNPLIDVSRNKKALHGNFQCTPIGVSMQNTIRIALAAICGLMA 415
Db 361 POIEVIRSTKMIEREINSVNDNPLIDVSRNKKALHGNFQCTPIGVSMQNTIRIALAAICGLMA 420
QY 417 LMFQFSELVNDVFNNGLPSPNLSGGRNPSLDYGFKGEIIMASYSCELOFLANPVTNHVQ 476
Db 421 LMFQFSELVNDVFNNGLPSPNLSGGRNPSLDYGFKGEIIMASYSCELOFLANPVTNHVQ 480
QY 477 SABOHNDVNSLGLISARKTAEAVDILKLMSSYVALCOSIDLRLLENNKSTVKNVSQVAK 536
Db 481 SABOHNDVNSLGLISARKTAEAVDILKLMSSYVALCOSIDLRLLENNKSTVKNVSQVAK 540
QY 537 QVARKVLMGVNGLHPSPRCEKDLRYVDREYFAYIDVCGSTYPLMKIHOVLDHALNNG 596
Db 541 QVARKVLMGVNGLHPSPRCEKDLRYVDREYFAYIDVCGSTYPLMKIHOVLDHALNNG 600
QY 597 LNNETKNTNTSIFOKIATFEELKVLIPKEVEGVRIAYENDTJSPNRIKACRSYPLRFAV 656
Db 601 LNNETKNTNTSIFOKIATFEELKVLIPKEVEGVRIAYENDTJSPNRIKACRSYPLRFAV 660

RESULT 7
T03663
phenylalanine ammonia-lyase (EC 4.3.1.5) - common tobacco
C:Species: Nicotiana tabacum (common tobacco)
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 16-Jul-1999
C:Accession: T03663
R:Legg, L.; Kohfritsch, O.; Fritig, B.; Legrand, M.
Plant Physiol 106, 877-886, 1994
A:Title: Phenylalanine ammonia-lyase in tobacco. Molecular cloning and gene expression
A:Reference number: T03663; MUID:95125127; PMID:7824656
A:Accession: T03663
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-712 <PELV>
A:Cross-references: EMBL:X78263; NID:q633596; PION:CAA55075.1; PTD:q633597
A:Experimental source: cultivar Samsun NN; tissue-type leaves after infection by tobacco
A:Genetics:
A:Gene: PAL
C:Function:
A:Description: catalyzes the deamination of L-phenylalanine to trans-cinnamic acid
A:Note: strongly induced during the hypersensitive reaction to TMV or to a fungal pathogen
C:Superfamily: histidine ammonia-lyase
C:Keywords: ammonia-lyase; carbon-nitrogen lyase
F:198-200/Cross-link: 5-imidazolone (Ala-Gly) #status predicted
F:199/Modified site: dehydroalanine (Ser) #status predicted

Query Match 85.6%; Score 3118.5; DB 2; Length 712;
Best Local Similarity 86.0%; Pred. No. 5,2e-192;
Matches 602; Conservative 38; Mismatches 59; Indels 1; Gaps 1;

```

```

QY 13 ELCTK-DPLMNGVAEALTGSHLDEYKKVAEERKRPVKLGSEETLVSGVAGIAAANDSD 71
Db 13 DPCVKVPLMNMMAOSI,KCSHLDYKKVAEERKRPVKLGSEETLVAGVAAIAAANDSD 72
QY 72 TVKVELSEARAAYKASSDWVMSNMKGTDSYVTTGATATSHPTKPGALOKELIRL 131
Db 73 TVKVELSEARAAYKASSDWVMSNMKGTDSYVTTGATATSHPTKPGALOKELIRL 132
QY 132 NAGTFCGCTSHLPHSATRAAMIVKINTLQCYSGIRPELLEATKILNNITPCLPL 191
Db 133 NAGVGNGTSCITLDPGSTRAMLVKINTLQCYSGIRPELLEATKILNNITPCLPL 192
QY 192 RGTITAGSDLVPLSYTNGTLTGPRNSKAVGPTGEVLNAEKAFAAEVEGFELOPKETL 251
Db 193 RGTITAGSDLVPLSYTNGTLTGPRNSKAVGPTGEVLNAEKAFAAEVEGFELOPKETL 252
QY 252 ALVNGTAVGSMASVLEFANVLALESEVLSAIFAEVNGKPEPTDHLTHKLNHPGQTE 311
Db 253 ALVNGTAVGSMASVLEFANVLALESEVLSAIFAEVNGKPEPTDHLTHKLNHPGQTE 312
QY 312 AAAIMEYLLDGSYVYKAAQVHEMDPLQKQKVAALKTSPQWIGQVIVNISTTKMIR 371
Db 313 AAAIMEYLLDGSYVYKAAQVHEMDPLQKQKVAALKTSPQWIGQVIVNISTTKMIR 372
QY 372 EINSVNDPLIDVSRNKALHG:NFLGTPIGVSMNTPLAIAALGKMFAPSELYNDFYN 431
Db 373 EINSVNDPLIDVSRNKALHG:NFLGTPIGVSMNTPLAIAALGKMFAPSELYNDFYN 432
QY 432 NGDPSNLGGGRNPSLDYGEKGEIAMAAYSCSELOFLANPVTNHVSAEDHODVNSLGI 491
Db 433 NGDPSNLGGGRNPSLDYGEKGEIAMAAYSCSELOFLANPVTNHVSAEDHODVNSLGI 492
QY 492 SARKTAFAVDILKIMSTYIVAIKOSIDLRLHFNKSTYKNTVQSAKAVITMGVNGTL 551
Db 493 SARKTAFAVDILKIMSTYIVAIKOSIDLRLHFNKSTYKNTVQSAKAVITMGVNGTL 552
QY 552 HPSRCEKDLRYVDREYFAYIDVCSGTYPIMOKLROYLDHALNGETKNTSIF 611
Db 553 HPSRCEKDLRYVDREYFAYIDVCSGTYPIMOKLROYLDHALNGETKNTSIF 612
QY 612 OKITFEELKVLIPKEVEGRIVAYENDTISIPNRKARSPYIPFVFEELGRTGTTE 671
Db 613 OKITFEELKVLIPKEVEGRIVAYENDTISIPNRKARSPYIPFVFEELGRTGTTE 672
QY 672 KVTSPGEFDFVFTAMCKQOIIDPLKLCIOGMNDEPLPIC 711
Db 673 KVTSPGEFDFVFTAMCKQOIIDPLKLCIOGMNDEPLPIC 712

```

RESULT 8

J02265
 phenylalanine ammonia-lyase (EC 4.3.1.5) : western balsam poplar x cottonwood
 C:Species: *Populus trichocarpa* x *Populus deltoides* (western balsam poplar x cottonwood)
 C:Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 28-May-1999
 C:Accession: J02265
 R:Subramaniam, R.; Reinold, S.; Mollitor, E.K.; Douglas, C.J.
 A:Title: Structure, inheritance, and expression of hybrid poplar (*Populus trichocarpa* x
 A:Reference number: J02265, MID:94151434, PMID:8108506
 A:Accession: J02265
 A:Molecule type: mRNA
 A:Residues: 1-715 <SRV>
 A:Cross references: GB:11747; NID:9169453; PUDN:AAA3805.1; PUD:9169454
 A:Note: The authors translated the codon AAG for residue 331 as Glu and AAG for residue
 C:Comment: The enzyme plays a role in linking primary metabolism to phenylpropanoid meta
 C:Superfamily: histidine ammonia-lyase
 C:Keywords: ammonia-lyase; carbon-nitrogen lyase
 F:201/203/Cross-link: 5 imidazolinone (Ala Gly) #status predicted
 F:202/Modified site: dehydroalanine (Ser) #status predicted

Query Match 85.5%; Score 3116, FR 2; Length 715;
 Best Local Similarity 84.1%; Pred. No. 7, 5e-192;

```

Matches 599; Conservative 56; Mismatches 53; Indels 4; Gaps 3;
QY 2 ENCHNGVYVNEICT--KIPLMNGVAEALTGSHLDEYKKVAEERKRPVKLGSEETLVSG 59
Db 6 KNG-YONGSLPGLCVNDPLSMGVAAEFAMKQSHLDYKKVAEERKRPVKLGSEETLV 64
QY 60 UVAIAAANISLVYKVELSEARAAYKASSDWVMSNMKGTDSYVTTGATATSHPTKPG 119
Db 65 QVAST-AGHDGVYKVELSEARAAYKASSDWVMSNMKGTDSYVTTGATATSHPTKPG 123
QY 120 GVALOKELIRLPLNAGTFCGCTSHLPHSATRAAMIVKINTLQCYSGIRPELLEATK 179
Db 124 GVALOKELIRLPLNAGTFCGCTSHLPHSATRAAMIVKINTLQCYSGIRPELLEATK 183
QY 180 PLANNITPCLPLRGTITAGSDLVPLSYTNGTLTGPRNSKAVGPTGEVLNAEKAFA 239
Db 184 PLANNITPCLPLRGTITAGSDLVPLSYTNGTLTGPRNSKAVGPTGEVLNAEKAFA 243
QY 240 GGFELDPKRYALVNGTAVGSMASVLEFANVLALESEVLSAIFAEVNGKPEPTDHL 299
Db 244 GGFELDPKRYALVNGTAVGSMASVLEFANVLALESEVLSAIFAEVNGKPEPTDHL 303
QY 300 THKLKHPGQIEAAAMEYLLDGSYVYKAAQVHEMDPLQKQKVAALKTSPQWIGQV 359
Db 304 THKLKHPGQIEAAAMEYLLDGSYVYKAAQVHEMDPLQKQKVAALKTSPQWIGQV 363
QY 360 EVIKSSIKMIRKINSVNDPLIDVSRNKALHG:NFLGTPIGVSMNTPLAIAALGKMF 419
Db 364 EVIKSSIKMIRKINSVNDPLIDVSRNKALHG:NFLGTPIGVSMNTPLAIAALGKMF 423
QY 420 AGFSELVNDYVNMALPNSLSG3FNP3SLDYGEKGEIAMAAYSCSELOFLANPVTNHV 479
Db 424 AGFSELVNDYVNMALPNSLSG3FNP3SLDYGEKGEIAMAAYSCSELOFLANPVTNHV 483
QY 480 OHNDVNSLGIISARKTAFAVDILKIMSTYIVAIKOSIDLRLHFNKSTYKNTVQSA 539
Db 484 OHNDVNSLGIISARKTAFAVDILKIMSTYIVAIKOSIDLRLHFNKSTYKNTVQSA 543
QY 540 KKVITMGVNGTLHPSRCEKDLRYVDREYFAYIDVCSGTYPIMOKLROYLDHALN 599
Db 544 KKVITMGVNGTLHPSRCEKDLRYVDREYFAYIDVCSGTYPIMOKLROYLDHALN 603
QY 600 GETEKNTSIFOKITFEELKVLIPKEVEGRIVAYENDTISIPNRKARSPYIPFVFE 659
Db 604 GENKNSSTSVPOKIFAEFELKALIPKEVEGRIVAYENDTISIPNRKARSPYIPFV 663
QY 660 REELGKPHLGHKAVTSPGGEFDFVFTAMCKQOIIDPLKLCIOGMNDEPLPIC 711
Db 664 REELGKPHLGHKAVTSPGGEFDFVFTAMCKQOIIDPLKLCIOGMNDEPLPIC 715

```

RESULT 9

J01858
 phenylalanine ammonia-lyase (EC 4.3.1.5) : common tobacco
 C:Species: *Nicotiana glauca* (common tobacco)
 C:Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 20-Jun-2000
 C:Accession: J01858
 R:Taguchi, G.; Sharan, M.; Gunda, K.; Yonagisawa, K.; Shimozuka, M.; Hayashida, N.; O
 J. Plant biochem. and biotech. 7, 79-84, 1998
 A:Title: Effect of methyl jasmonate and elicitor on PAL gene expression in tobacco cu
 A:Reference number: Z14446
 A:Accession: J01858
 A:Status: preliminary, translated from GH/PMH/JDHU
 A:Molecule type: DNA
 A:Residues: 1-712 <TAG>
 A:Cross references: EMBL:AB061199; PUDN:BA22447.1
 A:Experimental source: strain Bright yellow; tissue type callus
 C:Genetics:
 A:Gene: palA
 A:Introns: 130/2
 C:Function:
 A:Description: catalyzes the decarboxylation of L-phenylalanine to trans cinnamic acid
 C:Superfamily: histidine ammonia-lyase

C:Keywords: ammonia-lyase; carbon-nitrogen lyase
 F:198-200/Cross-link: 5-imidazolinone (Ala-Gly) #status predicted
 F:199/Modified site: dehydrolalanine (Ser) #status predicted

Query Match 85.3%; Score 3109.5; DB 2; Length 712;
 Best Local Similarity 86.1%; Pred. No. 1,9e-191;
 Matches 603; Conservative 35; Mismatches 61; Indels 1; Gaps 1;

QY 13 ELCTK-DPLNMGVAEALTGSHLDEVKKVAEPKPVYKLGGLLTYSOVAATAANDSD 71
 DB 13 DPCMKVDPNMMKADSLKSHLDEVKKVAEPKPVYKLGGLLTYSOVAATAANDSD 72
 QY 72 TVKVELSEARAGVASSDWMVSMNKKCTDSYVTTGCGATSHRRTKOGALOKELRFL 131
 DB 73 TVKVELSEARAGVASSDWMVSMNKKCTDSYVTTGCGATSHRRTKOGALOKELRFL 132
 QY 132 NGITGNGTSHITPHSATRAMIVRINTLLQYSGIRFELTALTKFLNNITPCLPL 191
 DB 133 NGVGVNGTSHITPHSATRAMIVRINTLLQYSGIRFELTALTKFLNNITPCLPL 192
 QY 192 RGTITASDPLVPLSYIAGLTGKPNKAVGPTGEVLNAEKFAAGVGGFFELQPEGL 251
 DB 193 RGTITASDPLVPLSYIAGLTGKPNKAVGPTGEVLNAEKFAAGVGGFFELQPEGL 252
 QY 252 ALVNGTAVGSMASVLPDANVIALISEVLSAIFAEVMQCKPEFTDHLTKLKHHPQIE 311
 DB 253 ALVNGTAVGSMASVLPDANVIALISEVLSAIFAEVMQCKPEFTDHLTKLKHHPQIE 312
 QY 312 AAIMEYITLDSGYKAAOKVHEMDPLQPKODRYALRTSPOMGPOIEVRSSTKMER 371
 DB 313 AAIMEYITLDSGYKAAOKVHEMDPLQPKODRYALRTSPOMGPOIEVRSSTKMER 372
 QY 372 EINSVNDNPLIDVSRNKKALHGGNFGPTIGVSMONTBLAIAIGKLMFAQFSELVNDYFN 431
 DB 373 EINSVNDNPLIDVSRNKKALHGGNFGPTIGVSMONTBLAIAIGKLMFAQFSELVNDYFN 432
 QY 432 NGLPNSLSCGRNPSIDYGFQKGETAMASGSELOFLANPVYNHVSQARHNOVDNSIGLI 491
 DB 433 NGLPNSLSCGRNPSIDYGFQKGETAMASGSELOFLANPVYNHVSQARHNOVDNSIGLI 492
 QY 492 SARKTAEAVDILKMSSTYLVALCVSIDLRHLEENKMSIVANTVSQVAKKVLTVGNGEL 551
 DB 493 SARKTAEAVDILKMSSTYLVALCVSIDLRHLEENKMSIVANTVSQVAKKVLTVGNGEL 552
 QY 552 HPSRCEKDLRVNDREYFAYIDVCGTTPLMOKLROUVLVDHALNNGETKNTNTSIF 611
 DB 553 HPSRCEKDLRVNDREYFAYIDVCGTTPLMOKLROUVLVDHALNNGETKNTNTSIF 612
 QY 612 OKIATFEELKAVLPRKVEGVRIAYENDTSLIPNRIKACRSYPLRYRVRPELGGPLTGE 671
 DB 613 OKIATFEELKAVLPRKVEGVRIAYENDTSLIPNRIKACRSYPLRYRVRPELGGPLTGE 672
 QY 672 KVTSPGEEDRVFTAMCKGQIIDLPLECLGSGNNGEPLPIC 711
 DB 673 KVTSPGEEDRVFTAMCKGQIIDLPLECLGSGNNGEPLPIC 712

RESULT 10

JC5872
 phenylalanine ammonia-lyase (EC 4.3.1.5) 1 - Lithospermum erythrorhizon
 C:Species: Lithospermum erythrorhizon
 C:Date: 09-Mar-1998 #sequence_revision 09-Mar-1998 #text_change 20-Jun-2000
 C:Accession: JC5872; PC4506
 R:Yazaki, K.; Kataoka, M.; Honda, G.; Severin, K.; Heide, L.
 Biosci. Biotechnol. Biochem. 61, 1995-2003, 1997
 A:Title: cDNA cloning and gene expression of phenylalanine ammonia-lyase in Lithospermum
 A:Reference number: JC5872, MUID 98101967, PMID:9418986
 A:Accession: JC5872
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-710 <YAL>
 A:Cross-references: DBJ:DB83075, NID g2911123; P10N:AAA24928 1; P10:g2911122
 A:Accession: PC4506

A:Molecule type: protein
 A:Residues: 348-479 -YAL2-
 C:Comment: This enzyme catalyzes the first committed step in multibranched phenylprop
 s of roots to protect them against pathogens in the soil
 C:Superfamily: histidine ammonia-lyase
 C:Keywords: ammonia-lyase; carbon-nitrogen lyase
 F:196-198/Cross-link: 5-imidazolinone (Ala-Gly) #status predicted
 F:197/Modified site: dehydrolalanine (Ser) #status predicted

Query Match 84.5%; Score 3077.5; DB 2; Length 710;
 Best Local Similarity 85.5%; Pred. No. 2.2e-189;
 Matches 594; Conservative 53; Mismatches 59; Indels 5; Gaps 2;

QY 1 MENCNHNQVNVNEICIKPLNMGVAEALTGSHLDEVKKVAEPKPVYKLGGLTYSVG 60
 DB 5 VENG---NGKIMERCKMPLNMEVASESKSHLDEVKKVAEPKPVYLAACKTLLIGQ 61
 QY 61 VAGTAAANSTVYKVEI SEAPACVAKASDWMVSMNKKCTDSYVTTGCGATSHRRTKOG 120
 DB 62 VAGTAAANSTVYKVEI SEAPACVAKASDWMVSMNKKCTDSYVTTGCGATSHRRTKOG 119
 QY 121 GALLKELKRLNAGIFGNGTSHITPHSATRAMIVRINTLLQYSGIRFELTALTKF 180
 DB 120 GALLKELKRLNAGIFGNGTSHITPHSATRAMIVRINTLLQYSGIRFELTALTKF 179
 QY 181 INNNTTPLLPGITITASDPLVPLSYIAGLTGKPNKAVGPTGEVLNAEKFAAGVGGFFELQPEGL 240
 DB 180 INNNTTPLLPGITITASDPLVPLSYIAGLTGKPNKAVGPTGEVLNAEKFAAGVGGFFELQPEGL 239
 QY 241 GFPELQPEGLALVNGTAVGSMASVLPDANVIALISEVLSAIFAEVMQCKPEFTDHLT 300
 DB 240 GFPELQPEGLALVNGTAVGSMASVLPDANVIALISEVLSAIFAEVMQCKPEFTDHLT 299
 QY 301 HKLKHHPQIEFAAIMEYITLDSGYKAAOKVHEMDPLQPKODRYALRTSPOMGPOIE 360
 DB 300 HKLKHHPQIEFAAIMEYITLDSGYKAAOKVHEMDPLQPKODRYALRTSPOMGPOIE 359
 QY 361 VIRSSTKMERINSVNDNPLIDVSRNKKALHGGNFGPTIGVSMONTBLAIAIGKLMFA 420
 DB 360 VIRSSTKMERINSVNDNPLIDVSRNKKALHGGNFGPTIGVSMONTBLAIAIGKLMFA 419
 QY 421 QHSELVNFYNNGLPSNLSGGKNSLUDYGFQKGETAMASGSELOFLANPVYNHVSQAR 480
 DB 420 QHSELVNFYNNGLPSNLSGGKNSLUDYGFQKGETAMASGSELOFLANPVYNHVSQAR 479
 QY 481 HNOVDNSIGLIISARKTAEAVDILKMSSTYLVALCVSIDLRHLEENKMSIVANTVSQVAK 540
 DB 480 HNOVDNSIGLIISARKTAEAVDILKMSSTYLVALCVSIDLRHLEENKMSIVANTVSQVAK 539
 QY 541 KVLTVGNGELHPSRCEKDLRVNDREYFAYIDVCGTTPLMOKLROUVLVDHALNNG 600
 DB 540 KVLTVGNGELHPSRCEKDLRVNDREYFAYIDVCGTTPLMOKLROUVLVDHALNNG 599
 QY 601 ETEKNTNTSIPOKIATFEELKAVLPRKVEGVRIAYENDTSLIPNRIKACRSYPLRYR 660
 DB 600 ETEKNTNTSIPOKIATFEELKAVLPRKVEGVRIAYENDTSLIPNRIKACRSYPLRYR 659
 QY 661 EELGSGFLTGEKVTSPGEEEDRVFTAMCKGQIIDLPLECLGSGNNGEPLPIC 711
 DB 660 EELGSGFLTGEKVTSPGEEEDRVFTAMCKGQIIDLPLECLGSGNNGEPLPIC 710

RESULT 11

S25303
 phenylalanine ammonia-lyase (EC 4.3.1.5) - garden pea
 C:Species: Pisum sativum (garden pea)
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 20-Jun-2000
 C:Accession: S25303
 R:Kawamura, S.; Yamada, T.; Tanaka, Y.; Sriptraeritsak, P.; Kato, H.; Ichinose, Y.; Ka
 Plant Mol. Biol. 20, 167-170, 1992
 A:Title: Molecular cloning of phenylalanine ammonia-lyase cDNA from Pisum sativum.
 A:Reference number: S25303, MUID:92385763, PMID:1515609
 A:Accession: S25303

A: Molecule type: mRNA

A: Residues: 1-723 : KAW

A: Cross references: EMBL: D10001; NID: 9217979; FIDN: MAA00865.1; FID: 9217980

C: Superfamily: histidine ammonia-lyase

C: Keywords: ammonia-lyase, carbon nitrogen lyase

F: 209-211/cross-link: 5 imidazolinone (Ala-Gly) *status predicted

F: 210/Modified site: dehydroalanine (Ser) *status predicted

Query Match 84.2%; Score 3070; DB 2; Length 723;

Best local similarity 83.9%; Pred. No. 6, 86-189;

Matches 594; Conservative 57; Mismatches 51; Indels 6; Gaps 3.

```

QY 3 NGNINAVGVNELICIKDPLMNGVAALTSILDEVKKVAEPRKVPVKGSEETLVSVQA 62
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 21 MAANNMKKVVNS---ADPLMNGVAALTSILDEVKKVAEPRKVPVKGSEETLVSVQA 77
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 63 CIAAANDSPVYVELSEARAGVASSDWYMSNMKGTDSTGVTGPGATSHRTKOGCA 122
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 78 AIAAHDHG--VKVELSEARAGVASSDWYMSNMKGTDSTGVTGPGATSHRTKOGCA 135
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 123 LQKRLIPLINAGIFGNGTSTSHLPASATRAAMIVRINTLQGYSGIRPELLEATKFLIN 182
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 136 LQKRLIPLINAGIFGNGTSTSHLPATRAAMIVRINTLQGYSGIRPELLEATKFLIN 195
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 183 NNITPELPEKGTITASSGLVPLSYIAGLLTGRPNKRAVPTGEVLAKEAFMAAGVEGGE 242
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 196 NNITPELPEKGTITASSGLVPLSYIAGLLTGRPNKRAVPTGEVLAKEAFMAAGVEGGE 254
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 243 PELQIKRGLAIVNGTAVGSMASMVIFDANVALLSEVLSAIFAVYMOCKRPFTHLTHK 302
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 255 FELQIKRGLAIVNGTAVGSMASMVIFDANVALLSEVLSAIFAVYMOCKRPFTHLTHK 314
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 303 LKHHPRQIFAAALIMFYIIMQSNYVKAACQVHHMDPLQKRCQRYALRTPQMLGPQIEVY 362
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 315 LKHHPRQIFAAALIMFYIIMQSNYVKAACQVHHMDPLQKRCQRYALRTPQMLGPQIEVY 374
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 363 RSSTKMIKREINSVNDPLIDVSRKALHGNQGTPIGVSMNDTRLATIAIGKLMFAQF 422
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 375 RPSSTKMIKREINSVNDPLIDVSRKALHGNQGTPIGVSMNDTRLATIAIGKLMFAQF 434
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 423 SPVNVIPYNGCIIPSNISGCHNPSIDYCFKQCEIAMAASYCELOFLANVTVNIVGSAEON 482
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 435 SELVNDFTYNNGLPSNLSASRNSLDYGFGESEIAMAASYCELOFLANVTVNIVGSAEON 494
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 483 QDVNSLQILISARKTAAVADVLKIMSSSTYVALCOSIDLKHLLENKSTVKNITVSOVAKK 542
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 495 QDVNSLQILISARKTAAVADVLKIMSSSTYVALCOSIDLKHLLENKSTVKNITVSOVAKK 554
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 543 LTMGVNDELHPSRCEKDLRYVDREYFAYIDDVCSGTYPLMOKLROYLVDAHLNGET 602
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 555 LTMGVNDELHPSRCEKDLRYVDREYFAYIDDVCSGTYPLMOKLROYLVDAHLNGET 614
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 603 EKNNTNISTOKIATFEELKVLIPKFEVPCVRIAYENDTISINRIKACSYPIYFVFRHE 662
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 615 EKNLNTNISTOKIATFEELKVLIPKFEVPCVRIAYENDTISINRIKACSYPIYFVFRHE 674
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 663 LQKRLIPLINAGIFGNGTSTSHLPASATRAAMIVRINTLQGYSGIRPELLEATKFLIN 710
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 675 LQKRLIPLINAGIFGNGTSTSHLPASATRAAMIVRINTLQGYSGIRPELLEATKFLIN 722
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 12

S22991

Phenylalanine ammonia-lyase (EC 4.3.1.5) 1 - soybean

C: Species: Glycine max (soybean)

C: Date: 06-Jan-1994 *sequence_revision 10-Nov-1995 *text_change 16-Jul-1999

C: Accession: S22991

R: Frank, R. L.; Wodkin, D. O.

A: Title: Sequence and structure of a phenylalanine ammonia-lyase gene from Glycine max.

A: Reference number: S22991. MIM: 92190550. PMID: 1794682

A: Molecule type: DNA

A: Residues: 1-713 <FRA>

A: Cross references: EMBL: X52754; NID: 914676; FIDN: TAA7129.1; PID: 918377

C: Genes: PAL1

A: Introns: 131/2

C: Superfamily: histidine ammonia-lyase

C: Keywords: ammonia-lyase, carbon nitrogen lyase

F: 199-201/cross-link: 5-imidazolinone (Ala-Gly) *status predicted

F: 200/Modified site: dehydroalanine (Ser) *status predicted

Query Match 84.2%; Score 3068; DB 2; Length 713;

Best local similarity 82.6%; Pred. No. 9e-189;

Matches 592; Conservative 59; Mismatches 54; Indels 12; Gaps 4;

```

QY 1 MNGN-HVNGVNEIC-----KDPINMGVAALTSILDEVKKVAEPRKVPVKGSEETLVSVQA 53
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MNGN-HVNGVNEIC-----KDPINMGVAALTSILDEVKKVAEPRKVPVKGSEETLVSVQA 57
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 54 ETLTVSOVAGIAANDSDTYVELSEARAGVASSDWYMSNMKGTDSTGVTGPGATSHRTKOG 113
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 58 ETLTVSOVAGIAANDSDTYVELSEARAGVASSDWYMSNMKGTDSTGVTGPGATSHRTKOG 115
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 114 HRRTKOGALQKRLIPLINAGIFGNGTSTSHLPASATRAAMIVRINTLQGYSGIRPEI 173
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 116 HRRTKOGALQKRLIPLINAGIFGNGTSTSHLPATRAAMIVRINTLQGYSGIRPEI 175
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 174 LEATKFLINNTIPCLIPNGTITASGLVPLSYIAGLLTGRPNKRAVPTGEVLAKEAFMAAG 233
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 176 LEATKFLINNTIPCLIPNGTITASGLVPLSYIAGLLTGRPNKRAVPTGEVLAKEAFMAAG 235
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 234 AAGVFGFPELOPKRGLALVNGTAVGSMASMVIFDANVALLSEVLSAIFAVYMOCKRPF 293
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 236 ELASINSEHYELQPKKGLALVNGTAVGSMASMVIFDANVALLSEVLSAIFAVYMOCKRPF 295
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 294 EFTDHLTHKIKHHPRQIFAAALIMFYIIMQSNYVKAACQVHHMDPLQKRCQRYALRTPQ 353
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 296 EFTDHLTHKIKHHPRQIFAAALIMFYIIMQSNYVKAACQVHHMDPLQKRCQRYALRTPQ 355
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 354 WICPQEVIVISSTKMIKREINSVNDPLIDVSRKALHGNQGTPIGVSMNDTRLATIAIGK 413
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 356 WICPQEVIVISSTKMIKREINSVNDPLIDVSRKALHGNQGTPIGVSMNDTRLATIAIGK 415
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 414 IGLKMFQSESELVNDFTYNNGLPSNLSASRNSLDYGFGESEIAMAASYCELOFLANVTVN 473
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 416 IGLKMFQSESELVNDFTYNNGLPSNLSASRNSLDYGFGESEIAMAASYCELOFLANVTVN 475
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 474 HVOSARQNDVNSLQILISARKTAAVADVLKIMSSSTYVALCOSIDLKHLLENKSTVKN 533
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 476 HVOSARQNDVNSLQILISARKTAAVADVLKIMSSSTYVALCOSIDLKHLLENKSTVKN 535
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 534 TVSOVAKKVLITMGNDELHPSRCEKDLRYVDREYFAYIDDVCSGTYPLMOKLROYLV 593
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 536 TVSOVAKKVLITMGNDELHPSRCEKDLRYVDREYFAYIDDVCSGTYPLMOKLROYLV 595
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 594 DHALNNGTETKNTNISTOKIATFEELKVLIPKFEVPCVRIAYENDTISINRIKACSYPI 654
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 596 DHALNNGTETKNTNISTOKIATFEELKVLIPKFEVPCVRIAYENDTISINRIKACSYPI 656
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 654 LTRFVYHLLGQCHLISFQVTSFGDEHIVPTIAMCKQGLIDPELLEATKFLIN 710
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 656 LTRFVYHLLGQCHLISFQVTSFGDEHIVPTIAMCKQGLIDPELLEATKFLIN 712
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 13

A44133

Phenylalanine ammonia-lyase (EC 4.3.1.5) - tomato

C: Species: Lycopersicon esculentum (tomato)

C: Date: 10-Mar-1994 *sequence_revision 10-Mar-1994 *text_change 21-May-1999

C: Accession: A44133; S51076

R: Lee, S.W.; Kobb, J.; Nazari, R.N.

A: Title: Truncated phenylalanine ammonia-lyase expression in tomato (Lycopersicon esculentum)

A: Reference number: A44133. MIM: 92291052; PMID: 1601854

Db 679 TGLTGTGNTSPGECDEKLFSAACGCKITIDPLEECIGFWNCAPIPLIC 725

RESULT 15
S04127

phenylalanine ammonia-lyase (EC 4.3.1.5) class II - kidney bean

C:Species: Phaseolus vulgaris (kidney bean)

C:Date: 28-Feb-1990 #sequence revision 28-Feb-1990 #text change 21-May-1999

C:Accession: S04127

C:Clamier, C L.; Edwards, K.; Brown, M.; Liang, X.; Hillier, S. J.; Bujewicz, G. P.; Dixon, R.

Plant Mol. Biol. 12, 367-383, 1989

A:Title: Phenylalanine ammonia-lyase gene organization and structure.

A:Reference number: S04127

A:Accession: S04127

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-712 <CRA>

C:Gene: PAL2

C:Superfamily: histidine ammonia-lyase

C:Keywords: ammonia-lyase; carbon nitrogen lyase

F:198-200/Cross-link: 5 imidazoliumone (Ald-Gly) #status predicted

F:199/Modified site: dehydroalanine (Ser) #status predicted

Query Match 83.9%; Score 3058; DB 2; Length 712;

Best local similarity 84.3%; Pred. No. 3, 9e-188;

Matches 585; Conservative 60; Mismatches 47; Indels 2; Gaps 1;

QY 18 DPLAMGVAAEALGSHLDEYKKVAEERKRVKLGGETLTVSOVAGIAANDSTVVEL 77
Db 21 DPLMAAAEALGSHLDEYKKVAEERKRVKLGGETLTVSOVAGIAANDSTVVEL 78
QY 78 SEARACVAKASIMVWMSNMKGTDSYCVTTCATSHRRTKOGALOKELIFLNAGITG 137
Db 79 AENRACVAKASIMVWMSNMKGTDSYCVTTCATSHRRTKOGALOKELIFLNAGITG 138
QY 138 NGTETSHLPASATRAAMIYRINTLGGYSGIREFLEAITFELNNITPCLPLKGTITA 197
Db 139 NGTETSHLPASATRAAMIYRINTLGGYSGIREFLEAITFELNNITPCLPLKGTITA 198
QY 198 SGDIYPLSTYAGLTGTPPNKAVGPTGEVLNAEKAFNAGVGGFFELQPKGALVNGT 257
Db 199 SGDIYPLSTYAGLTGTPPNKAVGPTGEVLNAEKAFNAGVGGFFELQPKGALVNGT 258
QY 258 AVSGMASVPIFDANVALISEVLSAIFAVMKGKPEFTDHLTHKLKHHPCOIFAAATME 317
Db 259 AVSGMASVPIFDANVALISEVLSAIFAVMKGKPEFTDHLTHKLKHHPCOIFAAATME 318
QY 318 YILDGSDYVKAQVHEMDEPLQKPKQDRYALKTSQWLDGPULVIRKSTKMLEREINSYN 377
Db 319 YILDGSDYVKAQVHEMDEPLQKPKQDRYALKTSQWLDGPULVIRKSTKMLEREINSYN 378
QY 378 DNPIDYSRKKALHGKNGFUSTPLGVSMONTFLAIAALGKLMAGFSELYNDYNNGLPSN 437
Db 379 DNPIDYSRKKALHGKNGFUSTPLGVSMONTFLAIAALGKLMAGFSELYNDYNNGLPSN 438
QY 438 LSGCRNPSLDYGFKQCHIAASVCSHIOPLANFVTHNVOSAPOHNDVNSIGLISARKTA 497
Db 439 LSGCRNPSLDYGFKQCHIAASVCSHIOPLANFVTHNVOSAPOHNDVNSIGLISARKTA 498
QY 498 EAVDITLMSSTLYVALCGSIDLRHLEENKSTVKNTVSQVAKVLTMGVNGELIIPSRFC 557
Db 499 EAVDITLMSSTLYVALCGSIDLRHLEENKSTVKNTVSQVAKVLTMGVNGELIIPSRFC 558
QY 558 EKDLLRVVDEYFAYIDVCGSTYPLMOKLROVLDAHLNNGETEKNTNTSIFOKIATF 617
Db 559 EKALLKVEREYFAYIDVCGSTYPLMOKLROVLDAHLNNGETEKNTNTSIFOKIATF 618
QY 618 EEEELKVVLPKRVGQVRIAYVNDYTLSTIPNRIKACRSTPLVRFVWEELGQGLTGKVTSPG 677
Db 619 EEEELKVVLPKRVGQVRIAYVNDYTLSTIPNRIKACRSTPLVRFVWEELGQGLTGKVTSPG 678

QY 678 EEEUVRFTAMCKGJILPLEETLUSWNSPELPLIC 711
Db 679 EECUKVFSAMCCKITIDPLEECIGFWNCAPIPLIC 712

Search completed: March 29, 2003, 02:13:15
Job time : 51 secs

.

.

.

.

.

.

GenCore version 5.1.4.P5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: March 29, 2003, 00:34:05 : Search time 39 seconds
(without alignments)
756.145 Million cell updates/sec

Title: US-09-964-992A-1
Perfect score: 3644
Sequence: 1 MENCNHNVCVNVNHCIKDPL 11DDLEHCICGMMCEPLPIC 711

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3206	88.0	667	1	PALY_HELAN
2	3186.5	87.4	716	1	PALY_PORCI
3	3185.5	87.4	716	1	PALY_PORCI
4	3178.5	87.2	718	1	PALY_PORCI
5	3171.5	87.0	708	1	PALY_PORCI
6	3154.5	86.6	714	1	PALY_PORCI
7	3149	86.4	715	1	PALY_PORCI
8	3118.5	85.6	712	1	PALY_PORCI
9	3116	85.5	715	1	PALY_PORCI
10	3115	85.5	717	1	PALY_PORCI
11	3109.5	85.3	712	1	PALY_PORCI
12	3097.5	85.0	721	1	PALY_PORCI
13	3096	85.0	713	1	PALY_PORCI
14	3088	84.7	715	1	PALY_PORCI
15	3077.5	84.5	710	1	PALY_PORCI
16	3077	84.4	722	1	PALY_PORCI
17	3070	84.2	723	1	PALY_PORCI
18	3068.5	84.2	713	1	PALY_PORCI
19	3068	84.2	713	1	PALY_PORCI
20	3067.5	84.2	725	1	PALY_PORCI
21	3058	83.9	712	1	PALY_PORCI
22	3051	83.7	705	1	PALY_PORCI
23	3049	83.7	720	1	PALY_PORCI
24	3042	83.5	704	1	PALY_PORCI
25	3039.5	83.4	725	1	PALY_PORCI
26	3039	83.4	717	1	PALY_PORCI
27	3004	82.4	717	1	PALY_PORCI
28	2994	82.2	718	1	PALY_PORCI
29	2988	82.0	708	1	PALY_PORCI
30	2985.5	81.9	710	1	PALY_PORCI
31	2953.5	81.1	707	1	PALY_PORCI
32	2928.5	80.4	710	1	PALY_PORCI
33	2920.5	80.1	707	1	PALY_PORCI

ALIGNMENTS

Result ID	Query Match	Length	DB ID	Description
34	2909	79.8	682	1
35	2908.5	79.8	700	1
36	2730.5	74.9	703	1
37	2672.5	73.3	620	1
38	2632.5	72.2	710	1
39	2590	71.1	710	1
40	2585.5	71.0	695	1
41	2548	69.6	590	1
42	2506	68.8	571	1
43	2500	68.5	701	1
44	2373.5	65.1	754	1
45	2220	60.9	506	1

ALIGNMENTS

RESULT 1
ID PALY_HELAN
AC 004058:
DT 15-JUL-1998 (rel. 36, created)
DT 15-JUL-1998 (rel. 36, last sequence update)
DT 15-JUN-2002 (rel. 41, last annotation update)
DE Phenylalanine ammonia-lyase (EC 4.3.1.5).
GN PAL.
OS Helianthus annuus (Common sunflower).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
OC Helianthaceae; Helianthus.
OX NCBI_Taxid=4232;
RN [1]
RP SEQUENCE FROM N.A.
RA Mazyra F.A., Salles S., Dreyer J., Brackel-Dreyer P.,
RA Tourville D., Ledoit G.;
RI "Isolation of a complete cDNA from sunflower";
RI (in) Plant Gene Register PCR98-108.
CC -1- FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE
CC -1- FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE
CC -1- VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLPROPANE SKELETON.
CC -1- CATALYTIC ACTIVITY: L-phenylalanine + trans-cinnamate + NH(3).
CC -1- PATHWAY: Phenylpropanoid biosynthesis; first step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (probable).
CC -1- PTM: CONTAINS AN ACTIVE SITE 4-METHYLENE-IMIDAZOLE-5-ONE (MIO),
CC -1- WHICH IS FORMED AUTOCATALYTICALLY BY CIRCULIZING AND DEHYDRATION
CC -1- OF RESIDUES ALA-SER-GLY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PAL / HISTIDASE FAMILY.
CC
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announc/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: Y124C1, CAA73065.1, ..
CC
CC DR HISSP: P21310; IBBP.
CC DR InterPro: IPR001106, phc/HIS_NH3lyase.
CC DR Pfam: PF00221; PALY_1.
CC DR Trifams: T10P01226, phc-am-lyase; 1.
CC DR PROSITE: PS00488; PAL_HISTIDASE; 1.
CC KW Lyase; Phenylpropanoid metabolism; Multiquery family.
CC FT SITE 197 199 MODIFIED TO FORM 4-METHYLENE-IMIDAZOLE-5-ONE (MIO).
CC FT SITE 197 199 MODIFIED TO FORM 4-METHYLENE-IMIDAZOLE-5-ONE (MIO).
CC SO SQUONCH 667 AA; 72076 MW; PHD30590DRH042 CMC64;
CC
CC Query Match: 88.0%, Score 3206, DB 1, Length 667;
CC Best Local Similarity: 94.4%, Pred. No. 1.4e-19;
CC Matches 627; Conservative 13; Mismatches 24; Indels 0; Gaps 0;

```

Db 1 MNGGTHVNSANGFCIKDPLMNGVAAEALTSIHDEVKKMGEEFKKPVVKAQ3ETLIVSQ 60
Qy 61 VAGIAAANDSDTVKVELSLAARACAKASSDWMMSMKKCTDSTCVTTGFCATSHRRKOC 120
Db 61 VAGISAAAGGNNVKVELSEAAAGAKASSDWMMSMKKCTDSTCVTTGFCATSHRRKOC 120
Qy 121 GALOKELIRFLNAGIFGNGTETSHTLPHSATKAAMIVRINTLLQYSGISIREILEATKFE 180
Db 121 GALOKELIRFLNAGIFGNGTETSHTLPHSATKAAMIVRINTLLQYSGISIREILEATKFE 180
Qy 181 LNNNTTPECLPLRGTTTASDVLPSYIAGLLTGPNSSKAVGPTGEVLNAEKAFAGVEG 240
Db 181 LNNNTTPECLPLRGTTTASDVLPSYIAGLLTGPNSSKAVGPTGEVLNAEKAFAGVEG 240
Qy 241 GFEELOPKKGLAVNGTAVGSGMASVLEFANVALLSEVLSALFAEVMQKPEFTDHLT 300
Db 241 GFEELOPKKGLAVNGTAVGSGMASVLEFANVALLSEVLSALFAEVMQKPEFTDHLT 300
Qy 301 HKLKHHPGQIEAAAMEYILDSGDVYKAAQVHEMDPLQKPRDORVALRTSPQWLGPOIE 360
Db 301 HKLKHHPGQIEAAAMEYILDSGDVYKAAQVHEMDPLQKPRDORVALRTSPQWLGPOIE 360
Qy 361 VIRSSKMTIERINSVNDNPLIDVSRNKLIGNGFQGPFGVSDNTRRLAIAIGKLMFA 420
Db 361 VIRSSKMTIERINSVNDNPLIDVSRNKLIGNGFQGPFGVSDNTRRLAIAIGKLMFA 420
Qy 421 QFSELVNDYFNNGLPSNLSGCHNPFLIDYFGKCEIAMAASYCELOFLANPTNVQSAEQ 480
Db 421 QFSELVNDYFNNGLPSNLSGCHNPFLIDYFGKCEIAMAASYCELOFLANPTNVQSAEQ 480
Qy 481 HHQDVNSLGLISAKRTAEVNDILKMSSTYLVALCOSTDLRHEENKSTVKNVSOVAK 540
Db 481 HHQDVNSLGLISAKRTAEVNDILKMSSTYLVALCOSTDLRHEENKSTVKNVSOVAK 540
Qy 541 KYLIMGVNGELHPSRPFCEKDLFVVDPEYFAYIDDVCSGTPTLMOKLROVLVDHALNNG 600
Db 541 KYLIMGVNGELHPSRPFCEKDLFVVDPEYFAYIDDVCSGTPTLMOKLROVLVDHALNNG 600
Qy 601 ETEKMTNFSIFOKIATFPEELKVLPLKEVEGVRIAYENDTISIPNRIKACHSPFLYFVR 660
Db 601 ETEKMTNFSIFOKIATFPEELKVLPLKEVEGVRIAYENDTISIPNRIKACHSPFLYFVR 660
Qy 661 EELG 664
Db 661 EELG 664

```

RESULT 2

```

PAL_PETCR
ID PAL_PETCR STANDARD: PRT: 716 AA.
AC P2481:
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phenylalanine ammonia-lyase 1 (EC 4.3.1.5).
CN PAL.
OS Petroselinum crispum (parsley) (Petroselinum hortense).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Apiales; Apiaceae; Petroselinum.
OC NCBI_TaxID=4043;
CX 111
RP SEQUENCE FROM N.A.
RX MEDLINE=89356635; PubMed=2767049;
RA Lohs R., Dietrich A., Hahnbrock K., Schulz W.;
RA "A phenylalanine ammonia-lyase gene from parsley: structure,
RT regulation and identification of elicitor and light responsive
RT cis-acting elements".
RI EMO J. 8:1641-1648(1989).
RM 12)
RP SEQUENCE FROM N.A.
RA Kang X., Logemann E., Hahnbrock K.;

```

```

RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP ACTIVE SITE AND MUTAGENESIS.
RX MEDLINE=9432630; PubMed=8050576.
RA Schuster R., Reley J.;
RT *serine-202 is the putative precursor of the active site
RT dehydroalanine of phenylalanine ammonia lyase. Site-directed
RT mutagenesis studies on the enzyme from parsley (Petroselinum crispum
RT L.)".
RI FEBS Lett. 349:252-254(1994).
CC -I- FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE
CC FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE
CC VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLPROPANE SKELETON.
CC -I- CATALYTIC ACTIVITY: L-phenylalanine - trans-cinnamate + NH(3).
CC -I- PATHWAY: phenylpropanoid biosynthesis; first step.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (probable).
CC -I- PFM: CONTAINS AN ACTIVE SITE 4-METHYLDENE-IMIDAZOLE-5-ONE (MIO),
CC WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION
CC OF RESIDUES ALA-SEP-GLY (BY SIMILARITY)
CC -I- SIMILARITY: BELONGS TO THE PAL / HISTIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X15473; CAA33500.1; -
DR EMBL: X16772; CAA34715.1; -
DR EMBL: Y07654; CAA68938.1; -
DR PIR: S04463; S04463.
DR HSSP: P21310; 1BBP.
DR InterPro: IPR001106; Phe/His_NH3lyase.
DR Pfam: PF00221; PAL. 1.
DR TIGRFAMS: TIGR01226; pne-am_lyase; 1.
DR PROSITE: PS00488; PAL_HISTIDASE; 1.
KW Lyase; Phenylpropanoid metabolism; Multigene family.
FT SITE 202 204 MODIFIED TO FORM 4-METHYLDENE-IMIDAZOLE-
FT MUTAGEN 203 203 5-ONE (BY SIMILARITY).
FT FT 210 210 S->A: COMPLETE LOSS OF ACTIVITY.
FT MUTAGEN 716 AA: 77828 MW: 23082 DALTONS CIRC64.
SQ
Query Match 87.4%; Score 3186.5; DB 1; Length 716;
Best Local Similarity 86.2%; Pred. No. 2.4e-190;
Matches 620; Conservative 46; Mismatches 42; Indels 11; Gaps 4;

```

OY 353 QMIGPOLFVIRRSSTKMIEREINSVNDNPLIDVSRNKAJHGNFOCTPIGVSMNDTRIALA 412
 DB 358 QMIGPOLFVIRRSSTKMIEREINSVNDNPLIDVSRNKAJHGNFOCTPIGVSMNDTRIALA 417
 OY 413 AICKLMFAQFSELVNDYFNKNSLRNSLQYVSKHSEIAMAASVSEIOPIANVVT 472
 DB 418 AICKLMFAQFSELVNDYFNKNSLRNSLQYVSKHSEIAMAASVSEIOPIANVVT 477
 OY 473 NHVQSAEQUHODVNSLGLISARKTAENVILKMSSTYVALQOSIDIRHEENKSTVK 532
 DB 478 NHVQSAEQUHODVNSLGLISARKTAENVILKMSSTYVALQOSIDIRHEENKSTVK 537
 OY 533 NTVSQAQKAVITMGVNGELIPSHFCHKDLRYVDRHYVAYIDVSCGYPIPMOKIPOVI. 592
 DB 538 NTVSQAQKAVITMGVNGELIPSHFCHKDLRYVDRHYVAYIDVSCGYPIPMOKIPOVI. 597
 OY 593 VDHALLNGEFTKNTNISIFOKIATFEELKVLLEKVEGVRLAYENDLISIPNRIKACRS 652
 DB 598 VDHALLNGEFTKNTNISIFOKIATFEELKVLLEKVEGVRLAYENDLISIPNRIKACRS 657
 OY 653 YPLRYFREELSGFELTGKVTSPGEFRTVFTAMKIDILPILLEGJHNMGEPLPIC 711
 DB 658 YPLRYFREELSGFELTGKVTSPGEFRTVFTAMKIDILPILLEGJHNMGEPLPIC 716
 RESULT 3
 PAL3_PETCR STANDARD: PRT: 716 AA.
 AC P45728;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 15-JUN-1995 (Rel. 41, Last sequence update)
 DE Phenylalanine ammonia-lyase 2 (EC 4.3.1.5).
 GN PAL2.
 OS Petroselinum crispum (Parsley) (Petroselinum hortense).
 OC Eukaryota: Viridiplantae, Streptophyta: Embryophyta, Tracheophyta;
 OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots;
 OC Asterales: euasterids II: Apiales: Apiaceae: Petroselinum.
 OX NCBI_TaxID=4043;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95010141; PubMed=7925471;
 RA Appert C., Logemann R., Halbrock K., Schmid J., Amrhein N.;
 RT Structural and catalytic properties of the four phenylalanine
 RT ammonia-lyase isoenzymes from parsley (Petroselinum crispum Nym.);
 PL Eur J Biochem 225:491-499(1994).
 CC -1- FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE
 CC FIRST PRACTICE IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE
 CC VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLPROPANE SKELETON.
 CC -1- CATALYTIC ACTIVITY: L-phenylalanine - trans-cinnamate + NH(3).
 CC -1- PATHWAY: Phenylpropanoid biosynthesis; first step.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Prokaryote).
 CC -1- PTM: CONTAINS AN ACTIVE SITE 4-METHYLHISTIDINE-IMIDAZOLE-5-ONE (MIO),
 CC WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION
 CC OF RESIDUES ALA-SER-GLY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PAL / HISTIDASE FAMILY.
 CC
 CC This SWISS-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DB EMIL: X81158, CAA57056.1;
 DB HSSP: P21310, 188F;
 DB InterPro: IPR001106, Phc/His_NH3lyase.
 DB Pfam: PF00221, PAL_1;
 DB TIGRfam: TIGR01226, phe-am_lyase; 1;
 DB PROSITE: PS00481, PAL_HISTIDASE; 1;
 KW Lyase, Phenylpropanoid metabolism, Multigene family.
 FT SITE 202 204 MODIFIED TO FORM 4-METHYLHISTIDINE-IMIDAZOLE-5-ONE

FT 5-ONE (BY SIMILARITY).
 SQ SEQUENCE: 716 AA: 77964 MW: 17470DB881BHF66 CMC64;
 Query Match 87.4%; Score 3185.5; DB 1: Length 716;
 Best Local Similarity 86.2%; Freq No 2 No-100;
 Matches 620; Conservative 45; Mismatches 43; Indels 11; Gaps 4;
 OY 1 MEMN-----HNGVYVNLGKIK--DPLNMGYAALATGSHLDYKKKVAAPRYVYKIG 52
 DB 1 MEMNVAATNGVNGVNGKDFCKTEDEPLTWGTAALAMTSTHLEKKKVAAPRYVYKIG 60
 OY 53 GELHIVSQVAGIAAAMSDIVAVVELSFAAKAGVAKSSIMVMMSMKGIISYGVIGQCAT 112
 DB 61 GELHIVSQVAAI-SARDGCGVIVELISFAARAGVAKSSIMVMMSMKGIISYGVIGQCAT 119
 OY 113 SHRTKQGNALQKELTFLPNAAGTFNGTETSHTLPHSNTPAAMVINTLLQYSGITFE 172
 DB 120 SHRTKQGNALQKELTFLPNAAGTFNGTETSHTLPHSNTPAAMVINTLLQYSGITFE 177
 OY 173 ILKATIKPLNNITPCLPIKGTITASGDIYPISTYAGLITGCHPNSKAVGPIGCVINAKKA 232
 DB 178 ILKATIKPLNNITPCLPIKGTITASGDIYPISTYAGLITGCHPNSKAVGPIGCVINAKKA 237
 OY 233 PAAQVQCGFPFELQPKGLAVNGTAVQSGMASMVLFPANVIALSEVLSAIFAVVMCK 292
 DB 238 PAAQVQCGFPFELQPKGLAVNGTAVQSGMASMVLFPANVIALSEVLSAIFAVVMCK 297
 OY 293 PFTDILTIKRLKIHLPVUTIAAIMEYLLGYSYVRAAKVIEEMDPLQPKQVVALPSTP 352
 DB 298 PFTDILTIKRLKIHLPVUTIAAIMEYLLGYSYVRAAKVIEEMDPLQPKQVVALPSTP 357
 OY 353 QMIGPOLFVIRRSSTKMIEREINSVNDNPLIDVSRNKAJHGNFOCTPIGVSMNDTRIALA 412
 DB 358 QMIGPOLFVIRRSSTKMIEREINSVNDNPLIDVSRNKAJHGNFOCTPIGVSMNDTRIALA 417
 OY 413 AICKLMFAQFSELVNDYFNKNSLRNSLQYVSKHSEIAMAASVSEIOPIANVVT 472
 DB 418 AICKLMFAQFSELVNDYFNKNSLRNSLQYVSKHSEIAMAASVSEIOPIANVVT 477
 OY 473 NHVQSAEQUHODVNSLGLISARKTAENVILKMSSTYVALQOSIDIRHEENKSTVK 532
 DB 478 NHVQSAEQUHODVNSLGLISARKTAENVILKMSSTYVALQOSIDIRHEENKSTVK 537
 OY 533 NTVSQAQKAVITMGVNGELIPSHFCHKDLRYVDRHYVAYIDVSCGYPIPMOKIPOVI. 592
 DB 538 NTVSQAQKAVITMGVNGELIPSHFCHKDLRYVDRHYVAYIDVSCGYPIPMOKIPOVI. 597
 OY 593 VDHALLNGEFTKNTNISIFOKIATFEELKVLLEKVEGVRLAYENDLISIPNRIKACRS 652
 DB 598 VDHALLNGEFTKNTNISIFOKIATFEELKVLLEKVEGVRLAYENDLISIPNRIKACRS 657
 OY 653 YPLRYFREELSGFELTGKVTSPGEFRTVFTAMKIDILPILLEGJHNMGEPLPIC 711
 DB 658 YPLRYFREELSGFELTGKVTSPGEFRTVFTAMKIDILPILLEGJHNMGEPLPIC 716
 RESULT 4
 PAL3_PETCR STANDARD: PRT: 716 AA.
 AC P45729;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phenylalanine ammonia-lyase 3 (EC 4.3.1.5).
 GN PAL3.
 OS Petroselinum crispum (Parsley) (Petroselinum hortense).
 OC Eukaryota: Viridiplantae, Streptophyta: Embryophyta, Tracheophyta;
 OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots;
 OC Asterales: euasterids II: Apiales: Apiaceae: Petroselinum.
 OX NCBI_TaxID=4043;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95010141; PubMed=7925471;

RA Appert C., Logemann E., Hahlbrock K., Schmid J., Amrhein N.;
 RT "Structural and catalytic properties of the four phenylalanine
 RT ammonia-lyase isoenzymes from parsley (Petroselinum crispum Nym.)";
 RL Eur. J. Biochem. 225:491-499(1994).
 CC -1- FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE
 CC FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE
 CC VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLPROPANE SKELETON.
 CC -1- CATALYTIC ACTIVITY: L-phenylalanine = trans-cinnamate + NH(3).
 CC -1- PATHWAY: phenylpropanoid biosynthesis: first step.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- PFM: CONTAINS AN ACTIVE SITE 4-METHYLIDENE-IMIDAZOLE-5-ONE (MIO),
 CC WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION
 CC OF RESIDUES ALA-SER-GLY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PAL / HISTIDASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X81159; CA57057.1; -
 DR HSSP: P21310; 188F.
 DR InterPro: IPR001106; Phe/His_NH3lyase.
 DR Pfam: PF00221; PAL; 1.
 DR TIGRFAMS: TIGR01226; phe.am.lyase; 1.
 DR PROSITE: PS00488; PAL_HISTIDASE; 1.
 KM Lyase; Phenylpropanoid metabolism; Multigene family.
 FT SITE 204 206 MODIFIED TO FORM 4-METHYLIDENE-IMIDAZOLE-
 FT 5-ONE (BY SIMILARITY).
 SQ SEQUENCE 718 AA; 78165 MW; 17786451FFBDC35 CRC64;
 Query Match 87.2%; Score 3178.5; DB 1; Length 718;
 Best Local Similarity 87.1%; Pred. No. 7.7e-190;
 Matches 620; Conservative 46; Mismatches 41; Indels 5; Gaps 4;
 3 NGNHVGVNVELCIK-DPLMWGVAEALTGSHDEVKKVAEERKPVYKLGTELTYSQ 60
 9 NG-HANGNDICDLMKREPLMWGVAEALTGSHDEVKVAEERKPVYKLGTELTYSQ 67
 61 VAGIAANDSDIVKELSEARAGYKASSDVMSNMKGTSYGVTTGFGATSRIRKKG 120
 68 VAAISARDDSG-VKVELSEARAGYKASSDVMSNMKGTSYGVTTGFGATSRIRKKG 126
 121 GALOKELIRFLNAGIFGNCOTET-SHTLPHSATRAAMIVRITLLQGYSGIFELLEATIK 179
 127 GALOKELIRFLNAGIFGSGAENGNTLPHSATRAAMIVRITLLQGYSGIFELLEATIK 186
 180 PLANNITTCPLRGHTTASGDLVPLSYAGLLTGRPNKAVGPTGEVLAERKFAAGVE 239
 187 PLNNHTTCLPLRGHTTASGDLVPLSYAGLLTGRPNKAVGPTGEVLAERKFAAGVE 246
 240 GGEFFELQPEGLATVNGTAVSGMAWYLPDANVALLSEVLSAIFAEVMOGKREFTDHL 299
 247 GGEFFELQPEGLATVNGTAVSGMAWYLPDANVALLSEVLSAIFAEVMOGKREFTDHL 306
 300 TKLHNHPQGLEAAAMEYIIDGSDYVKAQKVHEMDPLQKPKODRYALRFSPOWLGPOI 359
 307 TKLHNHPQGLEAAAMEYIIDGSDYVKAQKVHEMDPLQKPKODRYALRFSPOWLGPOI 366
 360 EYIRSTKMIEREINSVMDNPLIDVSRNKALHGNFGQCTPIGVSMNTRLAIAIGKLMF 419
 367 EYIRSTKMIEREINSVMDNPLIDVSRNKALHGNFGQCTPIGVSMNTRLAIAIGKLMF 426
 420 AOFSELYNDEVNGLPSNLISGGRNSLDYGFKEGELIAMAASCSLEOTLANVNNHVSAAE 479
 427 AOFSELYNDEVNGLPSNLISGGRNSLDYGFKEGELIAMAASCSLEOTLANVNNHVSAAE 486
 480 QHNDVNSLIGLISARKTEAVDIIKLMSSTYLVALCOSIDIRLHEEMKSTVKTYSOVA 539
 487 QHNDVNSLIGLISARKTEAVDIIKLMSSTYLVALCOSIDIRLHEEMKSTVKTYSOVA 546

QY 540 KVLITMGVNGELHPSRFEKDLIRVDRREYFAIIDDVCSITPMLOKLROVYDHALNN 599
 DB 547 KVLITMGVNGELHPSRFEKDLIRVDRREYFAIIDDVCSATYPLMOKLRETLVEHALNN 606
 QY 600 GFEKNTSTISFOKATATEEELKVLIPKEVGVIAVENLTSPNIRKACRSYPTREY 659
 DB 607 GKKNELSTISFOKIAAEDEDEKMLLPREVETAAALIESGNPAIPNIRKCRSYPYKREY 666
 QY 660 REELRGELTGKVTSGPEFDRVFTAMCKQIIDPLLECIGNNGEPLPIC 711
 DB 667 RELSTEVLTGKRVSPGEFEKVTAMSKKEIIDPLECLESGNAPLIPC 718
 RESULT 5
 ID PAL1_DAUCA STANDARD; PRT; 708 AA.
 AC 023865;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phenylalanine ammonia-lyase 1 (EC 4.3.1.5).
 GN PAL1.
 OS Daucus carota (Carrot).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
 OC NCBI_TaxID=4039;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Kurodagosun;
 RX MEDLINE=97478956; PubMed=9337617;
 RA Takeda J., Ozeki Y., Yoshida K.;
 RT "Action spectrum for induction of promoter activity of phenylalanine
 RT ammonia-lyase gene by UV in carrot suspension cells.";
 RL Photochem. Photobiol. 66:464-470(1997).
 CC -1- FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE
 CC FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE
 CC VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLPROPANE SKELETON.
 CC -1- CATALYTIC ACTIVITY: L-phenylalanine = trans-cinnamate + NH(3).
 CC -1- PATHWAY: phenylpropanoid biosynthesis: first step.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- PFM: CONTAINS AN ACTIVE SITE 4-METHYLIDENE-IMIDAZOLE-5-ONE (MIO),
 CC WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION
 CC OF RESIDUES ALA-SER-GLY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PAL / HISTIDASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D85850; BAA23367.1; -
 DR HSSP: P21310; 188F.
 DR InterPro: IPR001106; Phe/His_NH3lyase.
 DR Pfam: PF00221; PAL; 1.
 DR TIGRFAMS: TIGR01226; phe.am.lyase; 1.
 DR PROSITE: PS00488; PAL_HISTIDASE; 1.
 KM Lyase; Phenylpropanoid metabolism; Multigene family.
 FT SITE 194 196 MODIFIED TO FORM 4-METHYLIDENE-IMIDAZOLE-
 FT 5-ONE (BY SIMILARITY).
 SQ SEQUENCE 708 AA; 76845 MW; 0610411373B680E3 CRC64;
 Query Match 87.0%; Score 3171.5; DB 1; Length 708;
 Best Local Similarity 87.2%; Pred. No. 2e-189;
 Matches 619; Conservative 41; Mismatches 43; Indels 7; Gaps 3;
 2 ENGNHVGVNVELCI-KOPLMWGVAEALTGSHDEVKKVAEERKPVYKLGTELTYSQ 60
 DB 4 ENKNVYVLG-NGLCMKQDPLMWGVAEALTGSHDEVKKVAEERKPVYKLGTELTYSQ 61

CC	1-1	SUBCELLULAR LOCATION: Cytoplasmic (Probable)
CC	1-1	PTM CONTAINS AN ACTIVE SITE 4-METHYLMIDAZOLE-5-ONE (MIO), WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION OF RESIDUOUS ALA-SER-GLY (4H SIMILARITY).
CC	-1	SIMILARITY: BELONGS TO THE PAL / HISTIDASE FAMILY.
CC	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL installation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/annouce/or_send_an_email_to_license@sib-sib.ch).
CC	CC	EMBL: D26596; PAA05643.1; -
DR	DR	HSSD: P21310; 188F.
DR	DR	InterPro: IPR001106; phe/HIS_NH3lyase.
DR	DR	Pfam: PF00221; PAL; 1.
DR	DR	TIGRPFAMS: TIGR01226; phe_am_lyase; 1.
DR	DR	PROSITE: PS00488; PAL_HISTIDASE; 1.
KW	FM	Lyase; Phenylpropanoid metabolism.
FT	FT	SITE 200 202 MODIFIED TO FORM 4 METHYLMIDAZOLE-5-ONE (4H SIMILARITY).
SV	SV	SEQUENCE 714 AA; 77751 MW; 5F0EC7BBD0AB9071 CRC64;
Query Match	86.6%	Score 3154.5; DB 1; Length 714;
Best Local Similarity	86.0%	Pred. No. 2,4e-188;
Matches 610;	Conservative 46;	Mismatches 50; Indels 3; Gaps 2;
QY	4	GNHY-NCVYNELCIRPLRMVGYAALGSHLDEYKKMAEFKKYVKKLGGLITYGYVA 62
DB	8	GNVGSQSGSGFGLKQPLNMVGYAALMGSHLEEYKMEEFKPVVRUGETLLISQVA 67
QY	63	GIAAANDSDYVKKELESEARAGVAKASSIMWVESNMKCTDSYGYTCGCA:SHHRTKQCGA 122
DB	68	AIAYRGSE--VAVELSESAREGVKASSIMWVESNMKCTDSYGYTCGCA:SHHRTKQCGA 125
QY	123	LQKELLIFELNAGIFNGTSTSHLPKSNATPAAMIVP:NTLLGYST:PFPIELAITKTLN 182
DB	126	LQKELLIFELNAGIFNGTSTSHLPKSNATPAAMIVP:NTLLGYST:PFPIELAITKTLN 185
QY	183	NNIPICPLPCGTTTASGDLVPLSTYAGLLTGPSPSKAVGPTGEVLNKKPFAAGVEGGEF 242
DB	186	NNITPCPLPLGTTTASGDLVPLSTYAGLLTGRHNSKAVGPTGELINKEAFRLAGEVGEF 245
QY	243	FELQPKRGALVNGTAVGSGMASNVLPDANYIAL:SHVLSAIFAEVWQCKPEPTDHLTHK 302
DB	246	FELQPKRGALVNGTAVGSGMASNVLPDANYIAL:SEVLSAIFAEVWQCKPEPTDHLTHK 305
QY	303	LKHHPGJIEAALIMETILID:SDTYVAAARVIVEMDLPKPKGEPALALTSY:ML:POIEVI 362
DB	306	LKHHPGJIEAALIMETILID:SDSSYVAAAKLPHMDLPKPKQODKHALTSY:ML:GPILEVI 365
QY	363	PSTFKMIEPEINSVNDNPLILEVSPKMAI:HCINFGYPT:EGVSMENTPLATAI:IKLMIAOT 422
DB	366	PSTFKSIETEPINSVNDNPL:INVSFPAKALDG:NP:PTET:IVS:ML:NTPLAVAS:IKLMIAOT 425
QY	423	SELYNIDPYYNNGILP:SN:SGGRN:SLDYGRKGGH:IMASTYCSHIOH:ANIYV:INHVASWQHN 482
DB	426	SELYNIDPYYNNGILP:SN:SGGRN:SLDYGRKGGH:IMASTYCSHIOH:ANIYV:INHVASWQHN 485
QY	483	QDVAVSLCLISARRTAAVADILKLMSTYLVLCOSIDIRLIEENMKSTYNTVSYQAVKV 542
DB	486	QDVAVSLCLISARRTAAVADILKLMSTYLVLCOAVIDIRLIEENIRMTYKSVISQAVKAV 545
QY	543	LTM:SVN:FELHPSPE:EKDLLPVADVEVYFAYIDOV:STQTP:LMQKLPQV:LVTHALNN:DET 602
DB	546	LTM:SVN:FELHPSPE:EKDLLPVADVEVYFAYIDDP:CAIYPLMOKDQV:LVTHALNN:DET 605
QY	603	EKNINNIS:LOK:IAFEEELKVLIPKVEVGV:IAAFENITLS:IPK:IAQCSYPL:YKRVHRE: 662
DB	606	EKNINNIS:FOK:IAFEEELKVLIPKVEVGV:IAAFENITLS:IPK:IAQCSYPL:YKRVHRE: 665

```

QY 663 LKHL:FKVTSPEEFDPVFTAM-KKJITDPLELCEGWNSEPTIC 711
DB 666 IOTELLTGKRVASPGEEFDKVTALCKGEMIDPLMDCKEMGAPLPIG 714

RESULT 7
PAL3_TOBAC STANDARD: PRT: 715 AA.
AC P25872:
DB 01-MAY-1992 (Rel. 22, Created)
DB 01-MAY-1992 (Rel. 22, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phenylalanine ammonia-lyase (EC 4.3.1.5).
UN IPAL OR PALB.
OS Nicotiana tabacum (common tobacco).
OC Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N. A.
RA Fukasawa-Akeda T.;
RL Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N. A.
RA STRAIN=cv. Bright yellow; TISSUE=Callus;
RA Taguchi G., Sharan M., Gonda K., Yanagisawa K., Shimozaka M.,
RA Hayashida N., Okazaki M.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE
CC FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE
CC VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLPROPANE SKELETON.
CC -1- CATALYTIC ACTIVITY: L-phenylalanine = trans-cinnamate + NH(3).
CC -1- PATHWAY: Phenylpropanoid biosynthesis; first step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- PTM: CONTAINS AN ACTIVE SITE 4-METHYLENE-IMIDAZOLE-5-ONE (MIO),
CC WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION
CC OF RESIDUES ALA-SER-GLY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PAL / HISTIDASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: M84466; AAA34122.1;
DR EMBL: AB008200; BAA22948.1;
DR HSP: P21310; 188F.
DR InterPro: IPR001106; Phe/His_NH3lyase.
DR Pfam: PF00221; PAL; 1.
DR TIGRFAMs: TIGR01226; phe_am_lyase; 1.
DR PROSITE: PS00488; PAL_HISTIDASE; 1.
KM Lyase; Phenylpropanoid metabolism; Multigene family.
FT SITE 201 203 MODIFIED TO FORM 4-METHYLENE-IMIDAZOLE-5-ONE (MIO).
FT SITE 201 203 5-ONE (BY SIMILARITY).
SQ SEQUENCE 715 AA: 7780 MW; 03CBAE8527394C62 CRC64;

Query Match 86.4%; Score 3149; DB 1; Length 715;
Best Local Similarity 85.3%; Pred. No. 5.2e-188;
Matches 610; Conservative 45; Mismatches 56; Indels 4; Gaps 3;

```

```

DB 121 TKNQALDKEKELIRPLNAVPSNGTETSTLPHSNTTPAAMLVINTLLGYTSIPELLEA 180
QY 177 ITRPLNNITPCLPLKCTITASGDLVPLSYIAGLLTRPNSKAVGPTGVLENAEKAPAA 236
DB 181 ITRLTNSNTPLPLPGTITASGDLVPLSYIAGLLTRPNSKAVGPTGVLENAEKAPAA 240
QY 237 GVEGPFPELUPKEGALVNGTANGSMAWLPANVALLSEVLSATFEVNOGPEET 296
DB 241 GVGSGPFPELUPKEGALVNGTANGSMAWLPANVALLSEVLSATFEVNOGPEET 300
QY 297 DDLTKLKHKKVQLEAAAMEYILDSDYKAAAKVHEMDPLKFKQDFALRTSFQWIG 356
DB 301 DDLTKLKHKKVQLEAAAMEYILDSDYKAAAKVHEMDPLKFKQDFALRTSFQWIG 360
QY 357 POIEVIRSTKMIEREINSVNDPLIDVSRKKALHGNPGCTPGVSMNTRIALAIK 416
DB 361 POIEVIRATKMIEREINSVNDPLIDVSRKKALHGNPGCTPGVSMNTRIALAIK 420
QY 417 LMEAFSELYNDEYNGSLPSNLGSRNPLDYGEGEIMASVCSLOPLANPTNHO 476
DB 421 LMEAFSELYNDEYNGSLPSNLGSRNPLDYGEGEIMASVCSLOPLANPTNHO 480
QY 477 SAEJHNDVNSLGLISAKRTAEAVDILKMSSTYVALVGSIDLPLEENKSTVNTVS 536
DB 481 SAEJHNDVNSLGLISAKRTAEAVDILKMSSTYVALVGSIDLPLEENKSTVNTVS 540
QY 537 QYAKVLTLMGVNGELHPRPCEKDLRVDPREYFAYITDVGSTPYLMOKIROYVYDA 596
DB 541 QYAKVLTLMGVNGELHPRPCEKDLRVDPREYFAYITDVGSTPYLMOKIROYVYDA 600
QY 597 LNNGETENKNTSIFOKATFEELKVLPRKEVGVAIVENDLSIPNKIKACRSYPL 656
DB 601 MNGSEKKNVNSIIPQKIGAREDELKAVLPKEVSAARAALESQNALPNITPCRSPYLY 660
QY 657 KVPPELHGFLTSKVTSPGEEDPVPFTAM-KKJITDPLELCEGWNSEPTIC 711
DB 661 IOTELLTGKRVASPGEEFDKVTALCKGEMIDPLMDCKEMGAPLPIG 715

RESULT 8
PAL3_TOBAC STANDARD: PRT: 712 AA.
ID PAL3_TOBAC
AC P45733:
DB 01-NOV-1995 (Rel. 32, Created)
DB 01-NOV-1995 (Rel. 32, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phenylalanine ammonia-lyase (EC 4.3.1.5).
OS Nicotiana tabacum (common tobacco).
OC Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N. A.
RA STRAIN=cv. Samsun NN; TISSUE=leaf;
RA MEDLINE=95125127; PubMed=7824656;
RA Pellegrini L., Kohfutsch O., Fritig B., Legend M.;
KM "Phenylalanine ammonia-lyase in tobacco. Molecular cloning and gene
RT expression during the hypersensitive reaction to tobacco mosaic virus
RT and the response to a fungal elicitor."
RL Plant Physiol. 106:877-886(1994).
CC -1- FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE
CC FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE
CC VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLPROPANE SKELETON.
CC -1- CATALYTIC ACTIVITY: L-phenylalanine = trans-cinnamate + NH(3).
CC -1- PATHWAY: Phenylpropanoid biosynthesis; first step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- INDUCTION: STRONGLY INDUCED DURING THE HYPERSENSITIVE REACTION TO
CC TMV OR TO A FUNGAL ELICITOR.
CC -1- PTM: CONTAINS AN ACTIVE SITE 4-METHYLENE-IMIDAZOLE-5-ONE (MIO),
CC WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION
CC OF RESIDUES ALA-SER-GLY (BY SIMILARITY).

```

CC -1- SIMILARITY: BELONGS TO THE PAL / HISTIDASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announc/>

CC or send an email to license@isb-sib.ch).

DR EMBL: X78269; CAA55075.1; .

DR HSSP: P21310; 1B8F.

DR InterPro: IPR001106; Phe/His_NH3lyase

DR Pfam: PF00221; PAL; 1.

DR TIGRGRAMS: TIGR01226; phe.am.lyase; 1.

DR PROSITE: PS00488; PAL_HISTIDASE; 1.

KM Lyase, Phenylpropanoid metabolism, Multigene family.

FT SITE 198 200 MODIFIED TO FORM 4-METHYLDENE-IMIDAZOLE-5-ONE (BY SIMILARITY).

FT SEQUENCE 712 AA: 77510 MW: 66585FA0A83244F0 CRC64:

Query Match: 85.6%, Score 3118.5, DB 1; Length 712;

Host Local Similarity 86.0%; Pred. No. 4e-186;

Matches 602; Conservative 38; Mismatches 59; Indels 1; Gaps 1;

QY 13 ELGK-DPLMGVAALATGSHLEDEYKKNVAFRRKVVYALGSETLVSOAGTAAANDSD 71

DB 13 DPCVKVPIINMEHMAADSLGSHLDHYKKNVAFRRKVVYALGSETLVSOAGTAAANDSD 72

QY 72 TVVEVLESEARACVAKASVWMSMKGTDSYVTTGATSHPTKQVATLQKELIFPL 131

DB 73 TVVEVLESEARACVAKASVWMSMKGTDSYVTTGATSHPTKQVATLQKELIFPL 132

QY 132 NAGIFGNGTSHLPHSATRAAMIVRNILQGYSGAPFELIHAITKLNINIPPCPL 191

DB 133 NAGVFGNGTSCITLPOSGTRAMLVRLNLQGYSGIFELDELATKLNINIPPCPL 192

QY 192 RGTITASGDLVPLSYLAELTTPNSKAVPTREVLNAEKAFAAVGEVGFELQPKREL 251

DB 193 RGTITASGDLVPLSYLAELTTPNSKAVPTREVLNAEKAFAAVGEVGFELQPKREL 252

QY 252 ALVNGTAVSGSAMSVLFLDANVALLSEVLSAFAVVMQKREFFDLTHIKKHIPGLE 311

DB 253 ALVNGTAVSGSAMSVLFLDANVALLSEVLSAFAVVMQKREFFDLTHIKKHIPGLE 312

QY 312 AAALIMEVILDGSDYVKAAGVHMDPIQKIKQIKYALKISQWMLGHOIFVIRSSTKMIR 371

DB 313 AAALIMEVILDGSDYVKAAGVHMDPIQKIKQIKYALKISQWMLGHOIFVIRSSTKMIR 372

QY 372 EINSVNDNPLIDVSRKKALHGNFNGQTPTGVSNQDNTRIATAITGLMFAQFSELYNDVYN 431

DB 373 EINSVNDNPLIDVSRKKALHGNFNGQTPTGVSNQDNTRIATAITGLMFAQFSELYNDVYN 432

QY 432 NGIPSNISGGSRNPILDYGRKGSSEIINAAVSCSELOFLANVTNINOSARQINDVNSLGI 491

DB 433 NGIPSNISGGSRNPILDYGRKGSSEIINAAVSCSELOFLANVTNINOSARQINDVNSLGI 492

QY 492 SAKRTAAAVN::KIMSTSTVVALQOSTDIAHLENNKSTVKKNIVSQVAKV::ITMGVNGH 551

DB 493 SAKRTAAAVN::KIMSTSTVVALQOSTDIAHLENNKSTVKKNIVSQVAKV::ITMGVNGH 552

QY 552 IPRSEFCEKELIARVDEVEYFAYTDVCSSTYPLMOKLROVLYDHALNNGETKNTSTIF 611

DB 553 IPRSEFCEKELIARVDEVEYFAYTDVCSSTYPLMOKLROVLYDHALNNGETKNTSTIF 612

QY 612 OKITAFEEELKLVPEVEGVATVENDTSLIPNRIKARSPVCLRFVEVEELGSPFTLGE 671

DB 613 OKITAFEEELKLVPEVEGVATVENDTSLIPNRIKARSPVCLRFVEVEELGSPFTLGE 672

QY 672 KVISPGHEHDPVFTAMCKGQIIMYLLKCLGQWNGRPPLPIC 711

DB 673 KVISPGHEHDPVFTAMCKGQIIMYLLKCLGQWNGRPPLPIC 712

RESULT 9

ID PALY_POPTR SIAMAKU; PRT: 715 AA.

AC 945730;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Phenylalanine ammonia-lyase (EC 4.3.1.5).

OS PAL.

ON Populus trichocarpa (Western balsam poplar).

OC Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta;

OC Spermatophyta, Magnoliophyta, eudicotyledons, core eudicots, Rosidae;

OC eurosids I: Malpighiales, Salicaceae, Populus

OX NCBI_TaxID=3694;

RN 111

RP SHEDENCE FROM N.A.

RC STRAIN-P. trichocarpa X P. deltoides: TISSUE=leaf;

FX MEDLINE:94151434; PubMed:4108506;

FA Substratum R. Reinhold S., Molitor E K., Douglas C.J.;

KT *Structure, inheritance, and expression of hybrid poplar (Populus trichocarpa x Populus deltoides) phenylalanine ammonia-lyase genes.*;

PL Plant Physiol. 102:71-83(1993).

CC -1- FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLPROPANE SKELETON.

CC -1- CATALYTIC ACTIVITY: L-phenylalanine -> trans-cinnamate + NH3.

CC -1- PATHWAY: Phenylpropanoid biosynthesis, first step.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).

CC -1- PTM: CONTAINS AN ACTIVE SITE 4-METHYLDENE-IMIDAZOLE-5-ONE (M70), WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION OF RESIDUES ALA-SER-GLY (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE PAL / HISTIDASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announc/>

CC or send an email to license@isb-sib.ch).

DR EMBL: L11747; AAA3805.1; .

DR HSSP: P21310; 1B8F.

DR InterPro: IPR001106; Phe/His_NH3lyase.

DR Pfam: PF00221; PAL; 1.

DR TIGRGRAMS: TIGR01226; phe.am.lyase; 1.

DR PROSITE: PS00488; PAL_HISTIDASE; 1.

KM Lyase, Phenylpropanoid metabolism, Multigene family.

FT SITE 201 203 MODIFIED TO FORM 4-METHYLDENE-IMIDAZOLE-5-ONE (BY SIMILARITY).

FT SEQUENCE 715 AA: 77918 MW: 5896837A6B8A8246 CRC64:

Query Match: 85.5%; Score 3116; DB 1; Length 715;

Host Local Similarity 84.1%; Pred. No. 5.8e-186;

Matches 599; Conservative 56; Mismatches 54; Indels 4; Gaps 3;

QY 2 ENCHGVNGVNEVLCI--KIPINMGVAALATGSHLEDEYKKNVAFRRKVVYALGSETLV 59

DB 6 KNV-YQNGSISLSL/VNDPRPLISGVAAEAMKSGSHLEDEYKKNVAFRRKVVYALGSETLV 64

QY 64 QVAGTAAANDSTPVKVLSEAPAVKASSVWMSMKGTDSYVTTGATSHPTKQ 119

DB 65 QVASI-AGHIOGVKVELSRAPGVKASSVWMSMKGTDSYVTTGATSHPTKQ 124

QY 120 GVALLELLEFELNAGISGNTETSHLPHSATRAAMIVRNILQGYSGIFELDELATK 179

DB 124 GVALLELLEFELNAGISGNTETSHLPHSATRAAMIVRNILQGYSGIFELDELATK 184

QY 180 PLNNNIPICLPLAGCTITASGDLVPLSYLAELTTPNSKAVPTREVLNAEKAFAAVGE 239

DB 184 PLNNNIPICLPLAGCTITASGDLVPLSYLAELTTPNSKAVPTREVLNAEKAFAAVGE 244

QY 240 GCFPELQKREGIALVNCVAGSGMAWVLPDANVIALISEVLSAIFAEVMCKPTFFDHL 299
 DB 244 SGFFELQKREGIALVNCVAGSGMAWVLPDANVIALISEVLSAIFAEVMCKPTFFDHL 303
 QY 300 THKLKHHPGQIEAAAMEYIILDSGVYKAAQKHEMDPQKODRYALRTSPQWLGPOI 359
 DB 304 THKLKHHPGQIEAAAMEYIILDSGVYKAAQKHEMDPQKODRYALRTSPQWLGPOI 363
 QY 360 EYIRSTKTEIEREINSVNDNPLIDVSRKALHSGNFGOSTPIGVSMNDTPLATAIGKLMF 419
 DB 364 EYIRSTKTEIEREINSVNDNPLIDVSRKALHSGNFGOSTPIGVSMNDTPLATAIGKLMF 423
 QY 420 AOFSELVNDVFNNGLPNSISGGRNPSIDYGFKGGRIAMASYCELOFLANPVYNHVSAAE 479
 DB 424 AOFSELVNDVFNNGLPNSISGGRNPSIDYGFKGGRIAMASYCELOFLANPVYNHVSAAE 483
 QY 480 QHNQVNSLGLISAKKTAEVDILKLMSSYLVALCOSIDLRLHEENKSKYKNTVSQVA 539
 DB 484 QHNQVNSLGLISAKKTAEVDILKLMSSYLVALCOSIDLRLHEENKSKYKNTVSQVA 543
 QY 540 KKVLTIMGVNGELHPSRCEKDLRYVDREYFAYIDVCSGTYPIMOKLRQVLDHALNN 599
 DB 544 KKVLTIMGVNGELHPSRCEKDLRYVDREYFAYIDVCSGTYPIMOKLRQVLDHALNN 603
 QY 600 GETEKNTNTSIFOKIATFEELKVLVLPKEVEGVRIAYHNDTUSIPNRKACRSYPIYKRV 659
 DB 604 GENENKFSYVQKLEAEFEELKALPKREVESARAAYDSGNSAIDNRIKRCRSYPIYKRV 663
 QY 660 REELRGELTGEKVTSPGEFDFVFTAMCKGQIIDPLLECLSGNGNPELPIC 711
 DB 664 REELRGELTGEKVTSPGEFDFVFTAMCKGQIIDPLLECLSGNGNPELPIC 715
 RESULT 10
 PA12_PROAV STANDARD: PRT: 717 AA.
 AC 064963:
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phenylalanine ammonia-lyase 1 (EC 4.3.1.5).
 CN PAL1.
 OS *Prunus avium* (Cherry).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eustosids I; Rosales; Rosaceae; Amygdaloideae; *Prunus*.
 OX NCBI_TaxID=42229;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Summit;
 RA Wietsma P.A., Wu Z.:
 RT "A full-length cDNA for phenylalanine ammonia-lyase cloned from ripe
 RT Sweet Cherry fruit (*Prunus avium*)."
 RL (In) Plant Gene Register PCR98-184.
 CC -1- FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE
 CC FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLANILINE OF A WIDE
 CC VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYPROPANE SKELETON.
 CC -1- CATALYTIC ACTIVITY: L-phenylalanine + trans-cinnamate + NH(3).
 CC -1- PATHWAY: Phenylpropanoid biosynthesis (Probable).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- PTM: CONTAINS AN ACTIVE SITE 4-METHYLDIENE-IMIDAZOLE-5-ONE (MIO),
 CC WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION
 CC OF RESIDUES ALA-SER-GLY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PAL / HISTIDASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL, outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/annoumce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: AF036948; AAC78457.1; -
 DR HSSP: P21310.188F.
 DR InterPro: IPR001106; Phe/His_NH3lyase.
 DR Pfam: PF00221; PAL; 1.
 DR TIGRFAMs: TIGR01226; phe_am_lyase; 1.
 DR ProSite: PS00488; PAL_HISTIDASE; 1.
 KW lysase; phenylpropanoid metabolism; Multigene family.
 FT SITE 203 205
 FT 5-ONE (BY SIMILARITY).
 SQ SEQUENCE 717 AA; 77999 MM; B84DF90FA08F60B3 CRC64:
 Query Match 85.5%; Score 3115; DB 1; Length 717;
 Best Local Similarity 84.7%; Pred. No. 6,7e-186;
 Matches 604; Conservative 50; Mismatches 53; Indels 6; Gaps 5;
 QY 2 ENGHNVCNVN--ELCI-KDPLNMGVAAEALTGSHLDEYKKMAEPRKPVYKLGTEITLV 58
 DB 8 QNG-HKNGSVLELPICIKKDPINMGVAAFTIKSHLDEYKPMAYEYKPVYKLGSHLTI 66
 QY 59 SQVAGIAAANDSDIVKVELSEARAGYKASSDWMYFPMNKGTSYGVYTGCGATSHRRTK 118
 DB 67 SQVAAI-ATHDS-VKVELSEARAGYKASSDWMYFPMNKGTSYGVYTGCGATSHRRTK 124
 QY 119 OGCAIQKELIRPLNAGIFGNGTETSHLPHSATRAAMIYRIINTLQYSGIRPELLEAIT 178
 DB 125 OGCAIQKELIRPLNAGYFGSTKESGHTLPHQATFPAAMLVPIINTLQYSGIRPELLEAIT 184
 QY 179 KPLNNITPCLPLRGITIASGDIVPLSTYIGLLTGPRNSKAVGPTGEVINAFAAFAAGV 238
 DB 185 KPLNNITPCLPLRGITIASGDIVPLSTYIGLLTGPRNSKAVGPTGEVINAFAAFAAGV 244
 QY 239 ECGFPELQKREGIALVNCVAGSGMAWVLPDANVIALISEVLSAIFAEVMCKPTFFDHL 298
 DB 245 SGFFELQKREGIALVNCVAGSGMAWVLPDANVIALISEVLSAIFAEVMCKPTFFDHL 304
 QY 299 LTHKLKHHPGQIEAAAMEYIILDSGVYKAAQKHEMDPQKODRYALRTSPQWLGPOI 358
 DB 305 LTHKLKHHPGQIEAAAMEYIILDSGVYKAAQKHEMDPQKODRYALRTSPQWLGPOI 364
 QY 359 IEVPSSTKMTIEPINSVNDNPLIDVSPNKALHSGNFGOSTPIGVSMNDTPLATAIGKLM 418
 DB 365 IEVPSSTKMTIEPINSVNDNPLIDVSPNKALHSGNFGOSTPIGVSMNDTPLATAIGKLM 424
 QY 419 FAOFSELVNDVFNNGLPNSISGGRNPSIDYGFKGGRIAMASYCELOFLANPVYNHVSAAE 478
 DB 425 FAOFSELVNDVFNNGLPNSISGGRNPSIDYGFKGGRIAMASYCELOFLANPVYNHVSAAE 484
 QY 479 FQHNQVNSLGLISAKKTAEVDILKLMSSYLVALCOSIDLRLHEENKSKYKNTVSQVA 538
 DB 485 FQHNQVNSLGLISAKKTAEVDILKLMSSYLVALCOSIDLRLHEENKSKYKNTVSQVA 544
 QY 539 AKKVLTMGVNGELHPSRCEKDLRYVDREYFAYIDVCSGTYPIMOKLRQVLDHALNN 598
 DB 545 AKKVLTMGVNGELHPSRCEKDLRYVDREYFAYIDVCSGTYPIMOKLRQVLDHALNN 604
 QY 599 NGETEKNTNTSIFOKIATFEELKVLVLPKEVEGVRIAYHNDTUSIPNRKACRSYPIYKRV 658
 DB 605 NGETEKNTNTSIFOKIATFEELKVLVLPKEVEGVRIAYHNDTUSIPNRKACRSYPIYKRV 664
 QY 659 VEEELRGELTGEKVTSPGEFDFVFTAMCKGQIIDPLLECLSGNGNPELPIC 711
 DB 665 VEEELRGELTGEKVTSPGEFDFVFTAMCKGQIIDPLLECLSGNGNPELPIC 717
 RESULT 11
 PA12_TOHAC STANDARD: PRT: 712 AA.
 AC P35513; O22114;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 15-JUN-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phenylalanine ammonia-lyase (EC 4.3.1.5).
 GN PALA.

OS Nicotiana glauca (Common tobacco)
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta, eudicotyledons; core eudicots;
 OC Asterales; easterids I; Solanales; Solanaceae; Nicotiana.
 NCBI_TaxID=4097;
 RX MEDLINE=94218401; PubMed=8165251;
 RA Nagai N., Kitouchi H., Shimozaka M., Okazaki M.;
 KT Cloning and sequencing of a full length cDNA coding for
 RL Phenylalanine ammonia-lyase from tobacco cell culture.*;
 RL Plant Physiol. 104:1091-1092(1994).
 RN (2)
 RP REVISIONS.
 RA Nagai N., Kitouchi H., Shimozaka M., Okazaki M.;
 RL Submitted (Oct-1997) to the EMBL/GenBank/DDBJ databases.
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Bright yellow; TISSUE=Callus;
 RA Taguchi G., Sharan M., Gonda K., Yanagisawa K., Shimozaka M.,
 RA Hayashida N., Okazaki M.;
 RL Submitted (Oct-1997) to the EMBL/GenBank/DDBJ databases.
 CC -1- FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE
 CC FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE
 CC VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLPROPANE SKELETON
 CC -1- CATALYTIC ACTIVITY: L-phenylalanine = trans-cinnamate + NH(3).
 CC -1- PATHWAY: phenylpropanoid biosynthesis, first step
 CC -1- SUBCELLULAR LOCATION: cytoplasmic (Probable).
 CC -1- PTM: CONTAINS AN ACTIVE SITE 4-METHYLIMINE-IMIDAZOLE-5-ONE (MIO),
 CC WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION
 CC OF RESIDUES ALA-SER-GLY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PAL / HISTIDASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.slb.ch/announc/>
 CC or send an email to license@slb.slb.ch).
 CC
 DR EMBL: D17467; BAA22963.1;
 DR EMBL: AB008199; BAA22947.1;
 DR HSSP: P21310; 1BRF;
 DR InterPro: IPR001106; Phe/His_NH3lyase.
 DR Pfam: PF00221; PAL; 1;
 DR TIGRFAMs: TIGR01226; Phe.am.lyase. 1;
 DR PROSITE: PS00488; PAL_HISTIDASE; 1.
 KM Lyase; Phenylpropanoid metabolism; Multigene family.
 FT SITE 198 200
 FT 5-ONE (BY SIMILARITY).
 FT MODIFIED TO FORM 4-METHYLIMINE-IMIDAZOLE-
 FT SITE
 SO SHOUNCE 712 AA; 77345 MW; E65F8A867FEAC41 CnC64;
 Query Match 85 %; Score 3109.5; DR: 1; Length 712;
 Best Local Similarity 86 %; Pred No 1 5a-195;
 Matches 603; Conservative 35; Mismatches 61; Indels 1; Gaps 1.
 QY 13 ELICIT-DPMVGAALVATGSHLDKVKVAFRRKPVVAGLGGTLTVSQVAGTAAANDSD 71
 Db 13 DFCMKVDPFLMEVAAADSLKSHLDEKVKVAFRRKPVVAGLGGTLTVSQVAGTAAANDSD 72
 QY 72 TVVVELSEARAVKSSKVMWESMKKGTDSYGVTTGGFATGHPPTKQVQALQKELTFPL 131
 Db 73 TVVVELSEARAVKSSKVMWESMKKGTDSYGVTTGGFATGHPPTKQVQALQKELTFPL 132
 QY 132 NAGIFNGTETETLPHSTRAAMIVRINTLLOGSGIFELLELTIFLNNTNPPCLP 191
 Db 133 NAGVFGNGTETETLPHSTRAAMIVRINTLLOGSGIFELLELTIFLNNTNPPCLP 192
 QY 192 KGTITASGLVPLSYAGLITGCPNSKAVCPGVINAKKAAACVRCPPETGPKPKCI 251
 Db 193 KGTITASGLVPLSYAGLITGCPNSKAVCPGVINAKKAAACVRCPPETGPKPKCI 252

QY 252 ALVNGTAVSGMAGSVLPDANVALLSEVLSAIFAEVWGKPEPTDILTKLKHHPDIE 311
 Db 253 ALVNGTAVSGMAGSVLPDANVALLSEVLSAIFAEVWGKPEPTDILTKLKHHPDIE 312
 QY 312 AAATMEYIILDGSDYVYAAKQVHHRMDPKPKQKDRVALKTSQWAGPQIEVRSSTKMER 371
 Db 313 AAATMEIILDGSSYVYAAKQVHHRMDPKPKQKDRVALKTSQWAGPQIEVRSSTKMER 372
 QY 372 EINSVNMNPLIIVSKKALHGCNFOCTPIGVSMNTKIALAATGMLAPQSELYNDPN 431
 Db 373 EINSVNMNPLIIVSKKALHGCNFOCTPIGVSMNTKIALAATGMLAPQSELYNDPN 432
 QY 432 NTEPSNLSGPNPSTLYGKFGSEFVMAVSGSELOFLANVTHVVSAGQMDVNSLTLL 491
 Db 433 NTEPSNLSGPNPSTLYGKFGSEFVMAVSGSELOFLANVTHVVSAGQMDVNSLTLL 492
 QY 492 SARKTAAVVDILKMSSTIYVACOSTIDRLHEFMKSTVKNVTSQVAKVYIMVNDCL 551
 Db 493 SARKTAAVVDILKMSSTIYVACOSTIDRLHEFMKSTVKNVTSQVAKVYIMVNDCL 552
 QY 552 HPSRPFCKDILKRVDDRYVYFATIDVSCGYPLQKRLKQVLYVHALNGETKKNINISF 611
 Db 553 HPSRPFCKDILKRVDDRYVYFATIDVSCGYPLQKRLKQVLYVHALNGETKKNINISF 612
 QY 612 QKATFEELKVLPRVEGVAVFVENDTLSTPNPTKACPSYPLPREVEELGREGELTGE 671
 Db 613 QKATFEELKVLPRVEGVAVFVENDTLSTPNPTKACPSYPLPREVEELGREGELTGE 672
 QY 672 KVTSPGTEPVYVTPANCKGQIIPVILPCLQGNNGPRPLPG 711
 Db 673 KVTSPGTEPVYVTPANCKGQIIPVILPCLQGNNGPRPLPG 712
 RESULT 12
 PAL5_LYCES STANDARD: PRT: 721 AA.
 AC P26600;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE phenylalanine ammonia-lyase (EC 4.3.1.5) (PAL)
 GN PAL5.
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asterales; easterids I; Solanales; Solanaceae; Solanum.
 NCBI_TaxID=4081;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Bonny Best;
 RX MEDLINE=92291052; PubMed=1601854;
 RA Lee S.-W., Kobb E.J., Nazari R.N.;
 RT Truncated phenylalanine lyase expression in tomato
 RT (Lycopersicon esculentum).
 BL J. Biol. Chem. 267:11824-11830(1992).
 CC -1- FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE
 CC FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE
 CC VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLPROPANE SKELETON.
 CC -1- CATALYTIC ACTIVITY: L-phenylalanine = trans-cinnamate + NH(3).
 CC -1- PATHWAY: phenylpropanoid biosynthesis, first step.
 CC -1- SUBCELLULAR LOCATION: cytoplasmic (Probable).
 CC -1- PTM: CONTAINS AN ACTIVE SITE 4-METHYLIMINE-IMIDAZOLE-5-ONE (MIO),
 CC WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION
 CC OF RESIDUES ALA-SER-GLY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PAL / HISTIDASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.slb.ch/announc/>
 CC or send an email to license@slb.slb.ch).

```

CC EMBL: M90692; AAA34176.1;
DR HSSP: P21310; 188F;
DR InterPro: IPR001106, Phe/His_NH3lyase.
DR Pfam: PF00221; PAL, 1.
DR TIGRfams: TIGR01226; phe_am_lyase; 1.
DR Prosite: PS00484; PAL_HISTIDASE; 1.
KW Lyase: Phenylpropanoid metabolism; Multigene family.
FT SITE 207 209 MODIFIED TO FORM 4-METHYLDIENE-IMIDAZOLE-5-ONE (BY SIMILARITY).
FT SEQUENCE 721 AA; 78495 MW; 3184772693D50E7 CRC64;

Query Match 85.0%; Score 3097.5; DB 1; Length 721;
Best Local Similarity 83.4%; Pred. No. 8.3e-185;
Matches 598; Conservative 50; Mismatches 62; Indels 7; Gaps 3;

OY 1 MENGHNHGVNVELC-----IKDPLNMGVAEALTGSHLDVKKMAVEPRPKVKGGETT 55
DB 6 VONG-HVNGEAMDLCCKSTINNDPLNMEAEELGSHLDVKKMAVEPRPKVKGGETT 64
OY 56 LTVSQVAGIA-AANDSDTVKVELSEARAGYKASSDWMESEMNKGTDSYGVTTGFGATSH 114
DB 65 LTVAGVASIANVDKNSGVKVELSPARAGYKASSDWMDMGKGTDSYGVTTGFGATSH 124
OY 115 RRTKGGAOKELIFLNLNAGFNGTETSHLPHSATRAAMIVRINTLLOGYSIREIL 174
DB 125 RRTKNGGAOKHLIKPLNAGVFGNTESSHLPHSATRAAMIVRINTLLOGYSIREIL 184
OY 175 EATKFLNNNTIPCLPLGTTTASGDVPLSYAGILTGFRNSKAVGPTGEVLAAEKAPA 234
DB 185 EATITLNNNTIPCLPLGTTTASGDVPLSYAGILTGFRNSKAVGPTGEVLAAEKAPK 244
OY 235 AACGEGPPEIOPKICLALVNGTAVGSMASVLPANVALLSEV.SAIPAEVMDCKPE 234
DB 245 VAGVTSGEPELOPKGIALVNGTAVGSMASVLPFENMLAVMESEV.SAIPAEVMDCKPE 304
OY 295 FTDHHTHKHKHPGQIEAALIMEYLLDGSBYVKAOKVHEMDPQPKQJODEVYALFTSPQW 354
DB 305 FTDYTLHKHKHPGQIEAALIMEYLLDGSBYVKAOKVHEMDPQPKQJODEVYALFTSPQW 364
OY 355 LGDQIEVIRSSTKMIREKINSVNDPLIDVSRNKAHOGNPGTPIGVSMNTRIALAA1 414
DB 365 LGDQIEVIRSSTKMIREKINSVNDPLIDVSRNKAHOGNPGTPIGVSMNTRIALAA1 424
OY 415 GLTMAOPSELYNDVYNNGLPSNLSCGRNPGLDYCFKQGLTAMASYGSELOFLANPTNH 474
DB 425 GLTMAOPSELYNDVYNNGLPSNLSCGRNPGLDYCGAELTAMASYGSELOFLANPTNH 484
OY 475 VQSAEONODVNSGLISARKTAEAVDILKLMSSYTLVALCOSIDRLHEENMKSTVKT 534
DB 485 VQSAEONODVNSGLISARKTAEAVDILKLMSSYTLVALCOSIDRLHEENMKSTVKT 544
OY 535 VQVAKKVLTMGVNGELHPSRPECEDLKVYDREYVAYIDDVCSGYTPLMOKIROYLD 594
DB 545 VQVAKKVLTMGVNGELHPSRPECEDLKVYDREYVAYIDDVCSGYTPLMOKIROYLD 604
OY 595 HALNGETEKNNTNSIPKIAFFEBELKVLDPKVPQVRIAYENDTISIPNRKACGSSP 654
DB 605 HALNGETEKNNTNSIPKIAFFEBELKVLDPKVPQVRIAYENDTISIPNRKACGSSP 664
OY 655 LYRFAEELGSGFELTGERVTSPEEFEDRFTAMCKGQIIDPLELCELGSMNEPLPTG 711
DB 665 LYRFAEELGSGFELTGERVTSPEEFEDRFTAMCKGQIIDPLELCELGSMNEPLPTG 721

RESULT 13
PALY_DIGLA
AC 023924: STANDARD: PRT; 713 AA.
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phenylalanine ammonia-lyase (EC 4.3.1.5).

```

```

OS Digitalis lanata (Foxglove).
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots:
OC Asterales: Asteridiales: Lamiales, Veronicaceae: Digitalis.
OX NCBI_TaxId=49450;
RN [1]
RP SEQUENCE FROM N.A.
RA Theoringer C.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC - FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE
CC FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE
CC VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLALANINE-PANE SKELETON
CC - CATALYTIC ACTIVITY: L-phenylalanine = trans-cinnamate + NH(3).
CC - PATHWAY: Phenylpropanoid biosynthesis; first step.
CC - SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC - PTM: CONTAINS AN ACTIVE SITE 4-METHYLDIENE-IMIDAZOLE-5-ONE (MIO),
CC WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION
CC OF RESIDUES ALA-SER-GLY (BY SIMILARITY).
CC - SIMILARITY: BELONGS TO THE PAL / HISTIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ002221; CAA05251.1;
DR HSSP: P21310; 188F;
DR InterPro: IPR001106, Phe/His_NH3lyase.
DR Pfam: PF00221; PAL; 1.
DR TIGRfams: TIGR01226; phe_am_lyase; 1.
DR Prosite: PS00484; PAL_HISTIDASE; 1.
KW Lyase: Phenylpropanoid metabolism; Multigene family.
FT SITE 198 200 MODIFIED TO FORM 4-METHYLDIENE-IMIDAZOLE-5-ONE (BY SIMILARITY).
FT SEQUENCE 713 AA; 77732 MW; 747C664C65BDC66A CRC64;

Query Match 85.0%; Score 3096; DB 1; Length 713;
Best Local Similarity 84.0%; Pred. No. 1e-184;
Matches 600; Conservative 50; Mismatches 56; Indels 8; Gaps 4;

OY 1 MENGHNHGVNVELC-----IKDPLNMGVAEALTGSHLDVKKMAVEPRPKVKGGETT 58
DB 5 VENGHGN--NGPCVKNDDPLNMVAAAEELGSHLDVKKMAVEPRPKVKGGETT 61
OY 59 VQVAGIAANDSDTVKVELSEARAGYKASSDWMESEMNKGTDSYGVTTGFGATSHRPTK 118
DB 62 SVVAALAAKDE--VAVGLAESSKAVGKASSDWMESEMNKGTDSYGVTTGFGATSHRPTK 119
OY 119 OGGAIOKELIFLNLNAGFNGTETSHLPHSATRAAMIVRINTLLOGYSIREILEATT 178
DB 120 OGGAIOKELIFLNLNAGFNGTETSHLPHSATRAAMIVRINTLLOGYSIREILEATT 179
OY 179 KFLNNNTIPCLPIKRTTASGDVPLSYAGILTGFRNSKAVGPTGEVLAAEKAPAFAV 238
DB 180 KFLNNNTIPCLPIKRTTASGDVPLSYAGILTGFRNSKAVGPTGEVLAAEKAPFLAQA 239
OY 239 EGG-FFELQPEGLALVNGTAVGSMASVLPANVALLSEV.SAIPAEVMDGKREFTD 297
DB 240 NSGLEFFELQPEGLALVNGTAVGSMASVLPANVALLSEV.SAIPAEVMDGKREFTD 299
OY 298 HLTHLKHHPGQIEAALIMEYLLDGSBYVKAOKVHEMDPQPKQJODEVYALFTSPQW 357
DB 300 HLTHLKHHPGQIEAALIMEYLLDGSBYVKAOKVHEMDPQPKQJODEVYALFTSPQW 359
OY 358 QIEVIRSSTKMIREKINSVNDPLIDVSRNKAHOGNPGTPIGVSMNTRIALAAIGKI 417
DB 360 QIEVIRSSTKMIREKINSVNDPLIDVSRNKAHOGNPGTPIGVSMNTRIALAAIGKI 419
OY 418 MFAQPSSELVNDVYNNGLPSNLSCGRNPGLDYCFKQGLTAMASYGSELOFLANPTNHVOS 477

```

Db 420 MFAQSHLVNDVFNNGIIPSNISQGNHPSIDVYGHKSHIAMASYCSHIOFLANPVINHYQS 479
 QY 478 AEQINDVNSLGLISARKRAEVDILKMSSTYVALCOSIDRLLENMSTVNTYSQ 537
 Db 480 AEQINDVNSLGLISARKRAEVDILKMSSTYVALCOSIDRLLENMSTVNTYSQ 539
 QY 538 VAKKVLITMGVNGELHSRCCEKDIILVNDREYFAYIDVCGTPIYMOKILQVIVDAL 597
 Db 540 VAKKVLITMGVNGELHSRCCEKDIILVNDREYFAYIDVCGTPIYMOKILQVIVDAL 599
 QY 598 NGCEKNTNTSTFQKTAIFEEELKVLKPEVEGVPIAYENDTSTPNFTRKRSYPLYP 657
 Db 600 KNGHNKKNASTSTFQKTAIFEEELKVLKPEVEGVPIAYENDTSTPNFTRKRSYPLYP 659
 QY 658 FVEBELGRGLTGEKVTSPGEFDRVTAMCGOIIDLELLELUS3MNGEPDLC 711
 Db 660 FVEBELGRGLTGEKVTSPGEFDRVTAMCGOIIDLELLELUS3MNGEPDLC 713

RESULT 14

PALY_STYHU

ID PALY_STYHU STANDARD: PRT: 715 AA

AC 845732:

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Phenylalanine ammonia-lyase (EC 4.3.1.5).

GN PALY1.1.

OS Stylosanthes humilis (Townsille stylo).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eucosids 1; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;

OC Stylosanthes.

OX NCBI_TaxID=35628;

RN [1]

RP SOURCE: FROM N.A.

RC STRAIN: cv. Paterson; TISSUE: Stem;

RX MEDLINE=95357421; PubMed=7630950;

RA Manners J.M., McIntyre C.L., Nourse J.P.;

RT "Cloning and sequence of a cDNA encoding phenylalanine ammonia-lyase

from the tropical forage legume Stylosanthes humilis.";

RL Plant Physiol. 108:1301-1302(1995).

CC -1- FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE

FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE

VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLPROPANE SKELETON.

CC -1- CATALYTIC ACTIVITY: L-phenylalanine + trans-cinnamate + NH(3).

CC -1- PATHWAY: Phenylpropanoid biosynthesis; first step.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).

CC -1- PTM: CONTAINS AN ACTIVE SITE 4-METHYLUMINE-IMIDAZOLE-5-ONE (MIO),

WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION

OF RESIDUES ALA-SER-GLY (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE PAL / HISTIDASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation

at the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>or send an email to license@sib-ch).

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

Query Match 84.7%; Score 3088; D3 1; Length 715;

Best local similarity 85.7%; Pred No. 3,26-184;
 Matches 595, Conservative 47, Mismatches 50, Indels 2; Gaps 2;

QY 18 EPDNLVVAEALITSHDEFEVKKVAAEFKPVYKLTSTELTVGVAAIQAANLSTVVEL 77
 Db 24 DPLNMAAAALKSGHSDVKKRHHVSRKPLVNLGGUILLISQVAALAAANGQVSVOL 81
 QY 78 SPAAKAVKASSVWVWVSNKKGTSTGYTTCGATSHPTKGGALQKFTPLINAGITPG 137
 Db 82 SEASRAKAVKASSVWVWVSNKKGTSTGYTTCGATSHPTKGGALQKFTPLINAGITPG 141
 QY 138 NGTEITSHLPHSAIHAAMIVKINILUGYSGIRPELLEAFKFLNNNTTPCPIYRGTTA 197
 Db 142 NGTEITNTLPHVTRAMAILVRIINTLUGYSGIRPELLEAFKFLNNNTTPCPIYRGTTA 201
 QY 198 SIDLVLPLSYASLITLTPNSKAVPTPEVLAHAKFAAAVGEVFEELDPEKGLAVNT 257
 Db 202 SCPIVPLSYAGLITGHPNSKAVGPHGHTNAKFAVQAAGISDPEFELQPEKGLAVNT 261
 QY 258 AVGSCMASVILPFAVIALISTEVLISAFVAVGQKPEFTDHLTKHHPGQIQAALIME 317
 Db 262 PVSGLASVYLFANILVAYSEVLISAFVAVGQKPEFTDHLTKHHPGQIQAALIME 321
 QY 318 YIIJGSDVYVAAQVHMDPLQKPKODRYALRTSPQWIGPQIEVIRKSTKMLFRLNSVN 377
 Db 322 HILDGSSYVAAKKLHLDPLQKPKODRYALRTSPQWIGPQIEVIRKSTKMLFRLNSVN 381
 QY 378 DNLPLDVSFKKALHGNFQYTPYGVGMDTPLAIALTKLMAQFSELYNDVYNNGLPSN 437
 Db 382 DNLPLDVSFKKALHGNFQYTPYGVGMDTPLAIALTKLMAQFSELYNDVYNNGLPSN 441
 QY 438 ISCGRNLSLPGYKGEIIMASVCSLEQFLANPVINHYQSARHODVNSGLISARKTA 497
 Db 442 LSARKNSLSDYGFKEITELMASVCSLEQFLANPVINHYQSARHODVNSGLISARKTN 501
 QY 498 PAVDILKIMSTYVALCOSIDRLLENMSTVNTYSQVAKKVIYMGVNGELHPSRC 557
 Db 502 PAVDILKIMSTYVALCOSIDRLLENMSTVNTYSQVAKKVIYMGVNGELHPSRC 561
 QY 558 EKDLIPVHVEYFAYIDVCGTPIYMOKILQVIVDALNN:ELEKNTNTSTFQKTAIF 617
 Db 562 EKDLIPVHVEYFAYIDVCGTPIYMOKILQVIVDALNN:ELEKNTNTSTFQKTAIF 621
 QY 618 FEEELKVLKPEVEGVPIAYENDTSTPNFTRKRSYPLYPFVEELGKQPLTGCATVSG 677
 Db 622 FEEELKVLKPEVEGVPIAYENDTSTPNFTRKRSYPLYPFVEELGKQPLTGCATVSG 681
 QY 678 FEEELKVLKPEVEGVPIAYENDTSTPNFTRKRSYPLYPFVEELGKQPLTGCATVSG 711
 Db 682 FEEELKVLKPEVEGVPIAYENDTSTPNFTRKRSYPLYPFVEELGKQPLTGCATVSG 715

RESULT 15

PALI_LITTER

ID PALI_LITTER STANDARD: PRT: 710 AA

AC 049835:

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Phenylalanine ammonia-lyase 1 (EC 4.3.1.5) (PAL-1).

OS Lithospermum erythrorhizon.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asteridae; euasterids I; euasterids I; Invertidae sedis; Boraginaceae;

OC Lithospermum.

OX NCBI_TaxID=34254;

RN [1]

RP SOURCE: FROM N.A.

RX MEDLINE=98101947; PubMed=9418980;

RA Yasaki K., Katsuka M., Honda G., Severin K., Heide L.;

RT "cDNA cloning and gene expression of phenylalanine ammonia-lyase in

Lithospermum erythrorhizon.";

RL Biosci. Biotechnol. Biochem. 61:1995-2003(1997).

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw mod-1

Run on: March 29, 2003, 01:54:51, Search time 63 Seconds
(without alignments)
2325.389 Million cell updates/sec

Title: US-09-964-992a-1
Perfect score 3644
Sequence: 1 MENCNHNVCVNEU7KDPFL IICPLEYLAHWNREPLPIC 711

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671560

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database:

1: SPREMBL_21:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_ornithine:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3644	100.0	711	08W2E4	08W2E4 lactuca sat
2	3195.5	87.7	716	09MAX1	09MAX1 catharantus
3	3148.5	86.4	730	09M567	09M567 rubus idaeu
4	3107.5	85.3	708	094EN0	094EN0 rehmannia q
5	3106	85.2	721	09XFX5	09XFX5 citrifus clm
6	3100.5	85.1	711	09AX15	09AX15 pharbitis n
7	3093	84.9	710	094C45	094C45 manihot esc
8	3090	84.8	716	094AT7	094AT7 agastache r
9	3084.5	84.6	712	094F89	094F89 manihot esc
10	3056.5	83.9	718	09XFX6	09XFX6 citrifus clm
11	3049	83.7	715	094J66	094J66 populus kit
12	3045.5	83.6	725	094AN1	094AN1 arabidopsis
13	3009	82.6	717	08RW4	08RW4 arabidopsis
14	2995	82.2	687	094F90	094F90 manihot esc
15	2820.5	77.4	655	09XG83	09XG83 vigna ungu
16	2671	73.3	619	090545	090545 cucumis mel

17	2645	72.6	618	09ES00	09ES00 dianthus ca
18	2496	64.5	703	08VX57	08VX57 cea mayas (m
19	2400	65.9	532	081647	081647 capsicum ch
20	2332.5	64.0	681	08R023	08R023 pinus sylve
21	2331.5	64.0	681	08R049	08R049 pinus sylve
22	2325.5	63.8	681	08SAS7	08SAS7 pinus sylve
23	2321.5	63.7	681	08SAS6	08SAS6 pinus sylve
24	2320	60.1	544	09E138	09E138 trifolium ar
25	2194	59.2	549	094876	094876 hortensia var
26	1948	53.5	475	094875	094875 hortensia var
27	1887	51.8	497	094869	094869 hortensia var
28	1855	50.9	396	08VWP6	08VWP6 beta vulgar
29	1737	47.7	494	09SWP8	09SWP8 beta vulgar
30	1693	46.5	362	09F017	09F017 rubus idaeu
31	1675	46.0	363	09F018	09F018 rubus idaeu
32	1616	44.3	356	09S0M2	09S0M2 prunus pers
33	1573	43.2	395	08S3W3	08S3W3 cucumis sat
34	1368	37.5	339	043211	043211 trifolium ar
35	1365	37.5	315	094F91	094F91 manihot esc
36	1355	37.2	319	080406	080406 vitis vinif
37	1292.5	35.5	436	091EP1	091EP1 betula vort
38	1288.5	35.4	435	09M301	09M301 betula vort
39	1179	32.4	285	091EH2	091EH2 junifans nig
40	1063	29.2	240	0940D8	0940D8 brassica ra
41	977	26.8	217	09M627	09M627 collera arab
42	844.5	23.2	740	093967	093967 amantia mus
43	833	22.9	173	082476	082476 manihot esc
44	832	22.8	178	041293	041293 sorghum bic
45	811	22.3	186	09M626	09M626 coffea arab

ALIGNMENTS

RESULT 1	08W2E4	PRELIMINARY:	PRT:	711 AA.
ID	08W2E4			
AC	08W2E4:			
DT	01-MAR-2002 (TEMBLREL: 20, Created)			
DT	01-MAR-2002 (TEMBLREL: 20, last sequence update)			
DT	01-JUN-2002 (TEMBLREL: 21, last annotation update)			
DE	Phenylalanine ammonia-lyase.			
OS	Lactuca sativa (garden lettuce).			
OC	Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta:			
OC	Spermatophyta, Magnoliophyta, eudicotyledons, core eudicots:			
OC	Asteridae, euasterids II, Asterales, Asterales, Lactuca:			
OX	NCBI_TaxID=4236;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
PA	Chompos-Vargas R., Nonogaki H., Suslow T., Salviell M.:			
RT	"Characterization of phenylalanine ammonia-lyase (PAL) gene in wounded lettuce leaf tissue."			
RI	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AF299330; ANL55242.1;			
DR	Uniprot: IPR001106; Phe/His_NH3lyase.			
DR	Plan: PFO0221; PAL, 1.			
DR	PROSITE: PS00488; PAL_HISTIDASE; UNKNOWN_1.			
DR	LYASE.			
SO	SEQUENCE	711 AA:	77363 MW:	3120AR2877C29513 CMC64;
Query Match	100.0%	Score 3644:	DB 10:	Length 711:
First Local Similarity	100.0%	Pred No 9	79	223:
Matches 711,	Conservative 0,	Mismatches 0,	Indels 0:	Gaps 0:
QY	1 MENCNHNVCVNEU7KDPFLNMCVAAEAATGSHLDVKKKMAEPKVPVKIGCTIVSU 60			091560 dianthus ca
Pb	1 MENCNHNVCVNEU7KDPFLNMCVAAEAATGSHLDVKKKMAEPKVPVKIGCTIVSU 60			08VX57 cea mayas (m
QY	61 VAGIAAANDSDIVKVISSAARAGYKASSDWMPSMNGCTHSVGTGTCGATSHPPKKG 120			081647 capsicum ch
Pb	61 VAGIAAANDSDIVKVISSAARAGYKASSDWMPSMNGCTHSVGTGTCGATSHPPKKG 120			08R023 pinus sylve
QY	61 VAGIAAANDSDIVKVISSAARAGYKASSDWMPSMNGCTHSVGTGTCGATSHPPKKG 120			08R049 pinus sylve
Pb	61 VAGIAAANDSDIVKVISSAARAGYKASSDWMPSMNGCTHSVGTGTCGATSHPPKKG 120			08SAS7 pinus sylve
QY	61 VAGIAAANDSDIVKVISSAARAGYKASSDWMPSMNGCTHSVGTGTCGATSHPPKKG 120			08SAS6 pinus sylve
Pb	61 VAGIAAANDSDIVKVISSAARAGYKASSDWMPSMNGCTHSVGTGTCGATSHPPKKG 120			09E138 trifolium ar
QY	61 VAGIAAANDSDIVKVISSAARAGYKASSDWMPSMNGCTHSVGTGTCGATSHPPKKG 120			094876 hortensia var
Pb	61 VAGIAAANDSDIVKVISSAARAGYKASSDWMPSMNGCTHSVGTGTCGATSHPPKKG 120			094875 hortensia var
QY	61 VAGIAAANDSDIVKVISSAARAGYKASSDWMPSMNGCTHSVGTGTCGATSHPPKKG 120			094869 hortensia var
Pb	61 VAGIAAANDSDIVKVISSAARAGYKASSDWMPSMNGCTHSVGTGTCGATSHPPKKG 120			08VWP6 beta vulgar
QY	61 VAGIAAANDSDIVKVISSAARAGYKASSDWMPSMNGCTHSVGTGTCGATSHPPKKG 120			09SWP8 beta vulgar
Pb	61 VAGIAAANDSDIVKVISSAARAGYKASSDWMPSMNGCTHSVGTGTCGATSHPPKKG 120			09F017 rubus idaeu
QY	61 VAGIAAANDSDIVKVISSAARAGYKASSDWMPSMNGCTHSVGTGTCGATSHPPKKG 120			09F018 rubus idaeu
Pb	61 VAGIAAANDSDIVKVISSAARAGYKASSDWMPSMNGCTHSVGTGTCGATSHPPKKG 120			09S0M2 prunus pers
QY	61 VAGIAAANDSDIVKVISSAARAGYKASSDWMPSMNGCTHSVGTGTCGATSHPPKKG 120			08S3W3 cucumis sat
Pb	61 VAGIAAANDSDIVKVISSAARAGYKASSDWMPSMNGCTHSVGTGTCGATSHPPKKG 120			043211 trifolium ar
QY	61 VAGIAAANDSDIVKVISSAARAGYKASSDWMPSMNGCTHSVGTGTCGATSHPPKKG 120			094F91 manihot esc
Pb	61 VAGIAAANDSDIVKVISSAARAGYKASSDWMPSMNGCTHSVGTGTCGATSHPPKKG 120			080406 vitis vinif
QY	61 VAGIAAANDSDIVKVISSAARAGYKASSDWMPSMNGCTHSVGTGTCGATSHPPKKG 120			091EP1 betula vort
Pb	61 VAGIAAANDSDIVKVISSAARAGYKASSDWMPSMNGCTHSVGTGTCGATSHPPKKG 120			09M301 betula vort
QY	61 VAGIAAANDSDIVKVISSAARAGYKASSDWMPSMNGCTHSVGTGTCGATSHPPKKG 120			091EH2 junifans nig
Pb	61 VAGIAAANDSDIVKVISSAARAGYKASSDWMPSMNGCTHSVGTGTCGATSHPPKKG 120			0940D8 brassica ra
QY	61 VAGIAAANDSDIVKVISSAARAGYKASSDWMPSMNGCTHSVGTGTCGATSHPPKKG 120			09M627 collera arab
Pb	61 VAGIAAANDSDIVKVISSAARAGYKASSDWMPSMNGCTHSVGTGTCGATSHPPKKG 120			093967 amantia mus
QY	61 VAGIAAANDSDIVKVISSAARAGYKASSDWMPSMNGCTHSVGTGTCGATSHPPKKG 120			082476 manihot esc
Pb	61 VAGIAAANDSDIVKVISSAARAGYKASSDWMPSMNGCTHSVGTGTCGATSHPPKKG 120			041293 sorghum bic
QY	61 VAGIAAANDSDIVKVISSAARAGYKASSDWMPSMNGCTHSVGTGTCGATSHPPKKG 120			09M626 coffea arab

```

DB 121 GAGGQELIIRFLNAGTFCGCTEHSHTLPHSATRAAMIVAINTLLOGYSGIRREILFATKFR 180
QY 181 LNNNTTPELPLRGTTASGLVPLSTYAGLLTIRHNSKAV;PT-EVLNAEKAFAAVGS 240
DB 181 LNNNTTPELPLRGTTASGLVPLSTYAGLLTIRHNSKAV;PT-EVLNAEKAFAAVGS 240
QY 241 GFEELQPKGLAVNGTAVGSGMASMVLEFANVIALLESEVISAIFAEVYMGKPEFTDHLT 300
DB 241 GFEELQPKGLAVNGTAVGSGMASMVLEFANVIALLESEVISAIFAEVYMGKPEFTDHLT 300
QY 301 HKLKHHPQGLEAAALMEVYILDCSDYVKAOKVHEMDPLQKPKQDRAVALRTSPQW;GPOIF 360
DB 301 HKLKHHPQGLEAAALMEVYILDCSDYVKAOKVHEMDPLQKPKQDRAVALRTSPQW;GPOIF 360
QY 361 VRSSTKMTIEREINSVNDNPLIDVSRNKALHCGNFQGTPIGVSMNDNRLAIAAIGKLMFA 420
DB 361 VRSSTKMTIEREINSVNDNPLIDVSRNKALHCGNFQGTPIGVSMNDNRLAIAAIGKLMFA 420
QY 421 QFSELVNFYNNGLPSNLISGGNPSLDYGFKGCEIAMAASYCELOFLANPVTNHVQSAEQ 480
DB 421 QFSELVNFYNNGLPSNLISGGNPSLDYGFKGCEIAMAASYCELOFLANPVTNHVQSAEQ 480
QY 481 HNOOVNSLGLISAKTAAVADILKLMSSYLVALCOSTIDLRHLENNKSTVKNVSOVAK 540
DB 481 HNOOVNSLGLISAKTAAVADILKLMSSYLVALCOSTIDLRHLENNKSTVKNVSOVAK 540
QY 541 KYLTGVNGLSLHPSRCEKDLIRVVDREYFAYIDDVCSGTPIPMOKIROVYDHALNNG 600
DB 541 KYLTGVNGLSLHPSRCEKDLIRVVDREYFAYIDDVCSGTPIPMOKIROVYDHALNNG 600
QY 601 ETEKNTNISFOKIAATFEELKVLPLKKEVGVRIAYENDTISIPNRIKACSPYLFYFVR 660
DB 601 ETEKNTNISFOKIAATFEELKVLPLKKEVGVRIAYENDTISIPNRIKACSPYLFYFVR 660
QY 661 EELGKPLTGEKVTSPGCEFDVFTAMCKGQIIDPLLECLCGMNGEPLPIC 711
DB 661 EELGKPLTGEKVTSPGCEFDVFTAMCKGQIIDPLLECLCGMNGEPLPIC 711

RESULT 2
Q9MAX1 PRELIMINARY; PRT: 716 AA.
AC Q9MAX1:
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Phenylalanine ammonia-lyase.
OS Catharanthus roseus (Rosy periwinkle) (Madagascar periwinkle).
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicotis:
OC Asteridae: euasterids I: Gentianales: Apocynaceae: Rauvolfioidae:
OC Vinceae: Catharanthus.
OX NCBI_TaxID=4058;
RN [1]
RP SEQUENCE FROM N.A.
RA Kiyota S., Sakao K., Kim J.:
RT "cDNA of phenylalanine ammonia-lyase from Catharanthus roseus.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AH042520; BAA95629.1;
DR HSSP; P21310; 188F.
DR InterPro; IPR001106; Phe/His_NH3lyase.
DR Pfam; PF00221; PAL; 1.
DR TIGRfam; TIGR01226; phe_am_lyase.1.
DR PROSITE; PS00488; PAL_HISTIDINE; 1.
SQ
SEQUENCE 716 AA: 78241 MW: 1F16D4DEB8683B0B CRC64;

Query Match 87.7%; Score 3195.5; nr 10; Length 716;
Best Local Similarity 86.9%; Pred No.2-be-194;
Matches 617; Conservative 43; Mismatches 49; Indels 1; Gaps 1;
2 ENGNGVNGVNEICIKDPIIIMGVAEALTGSHLDEYVKKMAEFRRPVKLCGETIIVSQV 61

```

```

DB 8 ENGNGVNEIDVFCIKDPLIIMGMAADSLKSHLDEYVKKMAEFRRPVKLCGETIIVSQV 67
QY 62 AGLIAAANDSTGVKVELSENAFAVYKASSDVMYSESNKSTLYSV;VTTATATSHRPTK;63 121
DB 68 AAL-AARDYNAVKVLESHDAVAGVAKASSDVMYSESNKSTLYSV;VTTATATSHRPTK;63 126
QY 122 ALQKELIRFLNAGTFCGCTEHSHTLPHSATRAAMIVAINTLLOGYSGIRREILFATKFR 181
DB 127 ALQKELIRFLNAGTFCGCTEHSHTLPHSATRAAMIVAINTLLOGYSGIRREILFATKFR 186
QY 182 NNNITPCLPLNGTITASGDLPLSYIAGLLTGRNSKAVPTGCVINAKKAFMAACVNG 241
DB 187 NNNITPCLPLNGTITASGDLPLSYIAGLLTGRNSKAVPTGCVINAKKAFMAACVNG 246
QY 242 FFEELQPKGLAVNGTAVGSGMASMVLEFANVIALLESEVISAIFAEVYMGKPEFTDHLT 301
DB 247 FFEELQPKGLAVNGTAVGSGMASMVLEFANVIALLESEVISAIFAEVYMGKPEFTDHLT 306
QY 302 KKLKHHPQGLEAAALMEVYILDCSDYVKAOKVHEMDPLQKPKQDRAVALRTSPQW;GPOIFV 361
DB 307 KKLKHHPQGLEAAALMEVYILDCSDYVKAOKVHEMDPLQKPKQDRAVALRTSPQW;GPOIFV 366
QY 362 VRSSTKMTIEREINSVNDNPLIDVSRNKALHCGNFQGTPIGVSMNDNRLAIAAIGKLMFA 421
DB 367 VRSSTKMTIEREINSVNDNPLIDVSRNKALHCGNFQGTPIGVSMNDNRLAIAAIGKLMFA 426
QY 422 FSELVNDYNNGLPSNLISGGNPSLDYGFKGCEIAMAASYCELOFLANPVTNHVQSAEQ 481
DB 427 FSELVNDYNNGLPSNLISGGNPSLDYGFKGCEIAMAASYCELOFLANPVTNHVQSAEQ 486
QY 482 NODVNSLGLISAKTAAVADILKLMSSYLVALCOSTIDLRHLENNKSTVKNVSOVAK 541
DB 487 NODVNSLGLISAKTAAVADILKLMSSYLVALCOSTIDLRHLENNKSTVKNVSOVAK 546
QY 542 VLTGVNGLSLHPSRCEKDLIRVVDREYFAYIDDVCSGTPIPMOKIROVYDHALNNG 601
DB 547 VLTGVNGLSLHPSRCEKDLIRVVDREYFAYIDDVCSGTPIPMOKIROVYDHALNNG 606
QY 602 TEKNTNISFOKIAATFEELKVLPLKKEVGVRIAYENDTISIPNRIKACSPYLFYFVR 661
DB 607 SEKNVNTSIPQKIAAFHDELKTLVLPKKEVSARITALENGNAPINRIKHCSPYLFYFVR 666
QY 662 EELGKPLTGEKVTSPGCEFDVFTAMCKGQIIDPLLECLCGMNGEPLPIC 711
DB 667 EELGKPLTGEKVTSPGCEFDVFTAMCKGQIIDPLLECLCGMNGEPLPIC 716

RESULT 3
Q9M567 PRELIMINARY; PRT: 730 AA.
AC Q9M567:
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Phenylalanine ammonia-lyase 2 (EC 4.3.1.5).
OS Rubus idaeus (Raspberry).
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicotis:
OC Eurosid 1: Rosales: Rosaceae: Rosoideae: Rubus.
OX NCBI_TaxID=32247;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21437948; PubMed=11553751;
RA Kumar A., Ellis B.E.:
RT "The phenylalanine ammonia-lyase gene family in raspberry: structure,
expression, and evolution.";
RL Plant Physiol. 127:230-239 (2001).
DR EMBL; AF237955; AAF0224.1;
DR HSSP; P21310; 188F.
DR InterPro; IPR001106; Phe/His_NH3lyase.
DR Pfam; PF00221; PAL; 1.

```

DR TIGRFAMS: TIGR01226; phe_am_lyase; 1.
DR PROSITE: PS00488; PAL_HISTIDASE. 1.
KW lysase.
SD SOURCE: 730 AA; 79422 MW; 994241750CE113 CnC64;

Query Match 86.4%; Score 3148.5; DR 10; Length 730;
Best Local Similarity 83.9%; Pred. No. 2,8e-191;
Matches 610; Conservative 48; Mismatches 50; Indels 19; Gaps 4;

QY 2 ENC-NHNVGVN-----ELCIK-----DPLNMGVAEALTGSHLDKVKMAAF 44
DB 6 QNHSHHDM:QUNGSLDQGLVITKTESIKTGYSDPLNMGVAEASMTGSHLDKVPMAVEY 65
QY 45 RKPVNKGFEETLVSOVAG:IAANDSDTYKVELSEAAAGVASSRWMESSNKKGTDSYG 104
DB 66 KKVYKIGCTTITISOVAAL--ANHDSGVKVELAFASAGVASSRWMDSNKKGTDSYG 123
QY 105 VTGFSATSHRPTKQ:ALQKELIRFLNA:FRN:TESHT:PHSATRAAMIVRINTLLQ 164
DB 124 VTGFSATSHRPTKQ:ALQKELIRFLNAGVLRNTESSHT:PHSATRAAMIVRINTLLQ 183
QY 165 GYSGIRREH:IAATKPLNNNITTCPLRGT:IASG:VPL:SIAG:LLGRNSKAVPTG 224
DB 184 GYSGIRREH:IAATKPLNNNITTCPLRGT:IASG:VPL:SIAG:LLGRNSKAVPTG 243
QY 225 EVLNAAKFAAAGV:GFELOPKREG:ALVNGTAV:SGVMAVLPDANVALLESEV:SAI 284
DB 244 FTLMAAFAAAGV:GFELOPKREG:ALVNGTAV:SGVMAVLPDANVALLESEV:SAI 303
QY 285 FAEMVQKPEFTDHLTHLKLHHPG:TEAAAMEY:LLDSGVYKAAQVHEMPLQKPKOD 344
DB 304 FAEMVQKPEFTDHLTHLKLHHPG:TEAAAMEY:LLDSGVYKAAQVHEMPLQKPKOD 363
QY 345 RVALRTPQW:GQ:HV:HSST:KMI:FR:INSVNDNPL:DVSNKKA:HGCHGCTP:IGVSM 404
DB 364 RVALRTPQW:GQ:HV:HSST:KMI:FR:INSVNDNPL:DVSNKKA:HGCHGCTP:IGVSM 423
QY 405 DNTRLAIAIGKLMFAQSELVNDYFNNG:PLPSNL:SGRNP:SLDYFGKGEI:IMASV:SEL 464
DB 424 DNTRLAIAIGKLMFAQSELVNDYFNNG:PLPSNL:SGRNP:SLDYFGKGEI:IMASV:SEL 483
QY 465 QFLANPNTNIVASQ:EOHNDVNSLGL:ISARKTAEAVDILKMSSTYVALACOSIDLRIE 524
DB 484 QFLANPNTNIVASQ:EOHNDVNSLGL:ISARKTAEAVDILKMSSTYVALACOSIDLRIE 543
QY 525 ENKSTYKNTVSQAKRVLTGCVNGELHPSRC:EKDIL:RVNDEVFAV:IDVCG:STP:PL 584
DB 544 ENKSTYKNTVSQAKRVLTGCVNGELHPSRC:EKDIL:RVNDEVFAV:IDVCG:STP:PL 603
QY 585 MOKLROVLVDHALNGETEKNTST:FOK:IAFEEELKVLLEKVEGVRIAYENDTSLIP 644
DB 604 MOKLROVLVDHALNGETEKNTST:FOK:IAFEEELKVLLEKVEGVRIAYENDTSLIP 663
QY 645 NRIKARPSVPLRYFVREDELSP:FLTGEKVTSPGEEEDRFETAMGGQIIDPLELLELGSAN 704
DB 664 NRIKARPSVPLRYFVREDELSP:FLTGEKVTSPGEEEDRFETAMGGQIIDPLELLELGSAN 723
QY 705 GEPPLPIC 711
DB 724 GEPPLPIC 730

RESULT 4
Q94FNO PRELIMINARY: PRT: 708 AA.
AC Q94FNO:
DT 01-DEC-2001 (TIGR01226, 19, Created)
DT 01-DEC-2001 (TIGR01226, 19, Last sequence update)
DT 01-JUN-2002 (TIGR01226, 21, Last annotation update)
DE Phenylalanine ammonia-lyase (EC 4.3.1.5).
GN PAL1.
OS Rhamania glutinosa.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Matricophyta; Endicolydus; Core eudicots;
OC Asteridae; Euphorbiales; Lamiales; Inverta; Sedist; Rhamania;
OX NCBI_TaxId=99300;

RP SEQUENCE FROM N.A.
RA STRAIN:CV KEMISAN.
RA Yun S.J., Lee B.K., Park M.R., Shin J.B., Chin J.-C.;

RT "Molecular characterization of phenylalanine ammonia lyase gene in
Rhamania glutinosa";
RL Submitted (JUN-2001) to the EMBL/Genbank/DDBJ databases.
IR EMBL: AF401636; AAK84225.1;

DR InterPro: IPR001105; Phe/His NHase.
DR InterPro: IPR003880; Prantase-attach.
DR Pfam: PF00221; PAL. 1.

DR TIGRFAMS: TIGR01226; phe_am_lyase; 1.
DR PROSITE: PS00488; PAL_HISTIDASE; UNKNOWN. 1.
DR Lysase; PS00012; PHOSPHOPANTETHEINE; UNKNOWN. 1.

SU SEQUENCE 708 AA; 76835 MW; 186F0816F18FE CnC64;

Query Match 85.3%; Score 3107.5; DR 10; Length 708;
Best Local Similarity 84.3%; Pred. No. 1.1e-186;
Matches 602; Conservative 51; Mismatches 52; Indels 9; Gaps 4;

QY 1 MGNHNVGVNDELCT---KDLNMGVAEALTGSHLDKVKMAAFKPKVKTGSETLI 57
DB 1 MGNHNVGVNDELCT---KDLNMGVAEALTGSHLDKVKMAAFKPKVKTGSETLI 57
QY 58 VSOVAG:IAANDSDTYKVELSEAAAGVASSRWMESSNKKGTDSYGVTGFSATSHRT 117
DB 58 VSOVAG:IAANDSDTYKVELSEAAAGVASSRWMESSNKKGTDSYGVTGFSATSHRT 115
QY 118 KOGALQKELIRFLNAG:ITGNGTFSHTL:PHSATRAAMIVRINTLLQ:GYSGIRREH:IAAI 177
DB 116 KOGALQKELIRFLNAG:ITGNGTFSHTL:PHSATRAAMIVRINTLLQ:GYSGIRREH:IAAI 175
QY 178 TKELNNITPCLPRTITASQ:ELVPLSY:IASG:VPL:SIAG:LLGRNSKAVPTG:EVN:FEAAAG 237
DB 176 TKELNNITPCLPRTITASQ:ELVPLSY:IASG:VPL:SIAG:LLGRNSKAVPTG:EVN:FEAAAG 235
QY 238 VEGOFPELOKBEALVNG:IAVSGMAVLPDANVAL:ISV:SAI:FAVYQKPEFTD 297
DB 236 V-EGFEL:PKREG:ALVNG:IAVSGMAVLPDANVAL:ISV:SAI:FAVYQKPEFTD 294
QY 298 HTHKLKHHHPG:TEAAAMEY:LLDSGVYKAAQVHEMPLQKPKODYAL:RTSPQW:IGP 357
DB 295 HTHKLKHHHPG:TEAAAMEY:LLDSGVYKAAQVHEMPLQKPKODYAL:RTSPQW:IGP 354
QY 358 QIEVIRSTME:PEL:INSVNDNPL:DVSNKKA:HGCHGCTP:IGVSMONTL:IAIGKL 417
DB 355 QIEVIRSTME:PEL:INSVNDNPL:DVSNKKA:HGCHGCTP:IGVSMONTL:IAIGKL 414
QY 418 MPAQSELVNDYFNNG:PLPSNL:SGRNP:SLDYFGKGEI:IMASV:SELQFLANPNTNIVAS 477
DB 415 MPAQSELVNDYFNNG:PLPSNL:SGRNP:SLDYFGKGEI:IMASV:SELQFLANPNTNIVAS 474
QY 478 APOHNDVNSLGL:ISARKTAEAVDILKMSSTYVALACOSIDLRIE:ENKSTYKNTVSQ 537
DB 475 APOHNDVNSLGL:ISARKTAEAVDILKMSSTYVALACOSIDLRIE:ENKSTYKNTVSQ 534
QY 538 VAKKVLTMVNGELHPSRC:EKDIL:RVNDEVFAV:IDVCG:STP:PLMOKLROVLVDHAL 597
DB 535 VAKKVLTMVNGELHPSRC:EKDIL:RVNDEVFAV:IDVCG:STP:PLMOKLROVLVDHAL 594
QY 598 NNGTEKNTST:FOK:IAFEEELKVLLEKVEGVRIAYENDTSLIP:KEMISAN:PYK 657
DB 595 NNGTEKNTST:FOK:IAFEEELKVLLEKVEGVRIAYENDTSLIP:KEMISAN:PYK 654
QY 658 FVREDELSP:FLTGEKVTSPGEEEDRFETAMGGQIIDPLELLELGSAN:GEPPLPIC 711
DB 655 FVREDELSP:FLTGEKVTSPGEEEDRFETAMGGQIIDPLELLELGSAN:GEPPLPIC 708


```

QY 480 QHNDVNSLILISAKTIAEAVOLIKLMSSTYLVALUGSIDRHEENKMKSTYKNTVSQVA 539
DB 480 QHNDVNSLILISAKTIAEAVOLIKLMSSTYLVALUGSIDRHEENKMKSTYKNTVSQVA 539
QY 540 KKVLTMSVNGELSPRCECKDLRVDRREYFAYIDDVSCSTYPLMOKLROVLVDHAIANN 599
DB 540 KKVLTMSVNGELSPRCECKDLRVDRREYFAYIDDVSCSTYPLMOKLROVLVDHAIANN 599
QY 600 GEFTEKNTNISFOKIAFPEHEELKVLIPKRVBCVRIAYENDLISITNRIKACKSYPLNRPV 659
DB 600 GESEKNTNISFOKIAFPEHEELKVLIPKRVBCVRIAYENDLISITNRIKACKSYPLNRPV 659
QY 660 FEELSKFTTGEKVTSPSEEFQVFTAMCKSLIDPLELLELGGMMKEPPLPIC 711
DB 660 KEGICFTLITGEKVTSPSEEFQVFTAMCKSLIDPLELLELGGMMKEPPLPIC 711

RESULT 7
Q94C45 PRELIMINARY: PRT: 710 AA.
AC 094C45
DT 01-DEC-2001 (TREMblrel, 19, Created)
DT 01-DEC-2001 (TREMblrel, 19, Last sequence update)
DE 01-JUN-2002 (TREMblrel, 21, Last annotation update)
DE Phenylalanine ammonia-lyase 1 (EC 4.3.1.5).
CN PAL1.
OS Manihot esculenta (Cassava) (Manihot).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots; Rosidae;
OC eucots I; Malpighiales; Euphorbiaceae; Manihot.
OX NCBI_TaxID=3983;
RN 111
RP SHOUNCE FROM N.A.
RA Han Y., Boeching J.R.;
RT "Cassava phenylalanine ammonia-lyase expressed during post-harvest
RT physiological deterioration".
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY036011; AAK62030.1;
DR InterPro: IPR001106; Phc/Hls_NH3lyase.
DR Pfam: PF00221; PAL; 1.
DR TIGRfams: TIGR01226; Phc-am_lyase; 1.
DR PROSITE: PS00488; PAL_HISTIDASE; UNKNOWN; 1.
KW Lyase.
SQ SHOUNCE 710 AA: 77559 MW; F236AEH1PZHA1C5C CRC64.

Query Match 84.9%; Score 3093; DH 10; Length 710;
Best Local Similarity 85.3%; Pred. No. 9e-188;
Matches 592; Conservative 52; Mismatches 48; Indels 2; Gaps 1.

QY 18 DPLNMCVAAHAIQSHLDHVKKKVAFRRPVKLAGFETVVSQVAGIAAANDSDTVKYL 77
DB 18 DPLNMCVAAHAIQSHLDHVKKKVAFRRPVKLAGFETVVSQVAGIAAANDSDTVKYL 77
QY 19 DPLNMCVAAHAIQSHLDHVKKKVAFRRPVKLAGFETVVSQVAGIAAANDSDTVKYL 76
DB 19 DPLNMCVAAHAIQSHLDHVKKKVAFRRPVKLAGFETVVSQVAGIAAANDSDTVKYL 76
QY 78 SEARARAVKRSKSSQWVMEKMKGTSTSTGVTTFGATSHPTKQVSLQVELLIFLNAGIFG 137
DB 78 SEARARAVKRSKSSQWVMEKMKGTSTSTGVTTFGATSHPTKQVSLQVELLIFLNAGIFG 137
QY 77 SEARARAVKRSKSSQWVMEKMKGTSTSTGVTTFGATSHPTKQVSLQVELLIFLNAGIFG 136
DB 77 SEARARAVKRSKSSQWVMEKMKGTSTSTGVTTFGATSHPTKQVSLQVELLIFLNAGIFG 136
QY 138 NGFTSTHUPHSAFRAAMIVRIINTLLQYSGIFRFLFAITFLNNNTTTPCLPLGCTITA 197
DB 138 NGFTSTHUPHSAFRAAMIVRIINTLLQYSGIFRFLFAITFLNNNTTTPCLPLGCTITA 197
QY 137 NGQSCITLSTHTTRAAMLVRIITLLQYSGIFRFLFAITFLNNNTTTPCLPLGCTITA 196
DB 137 NGQSCITLSTHTTRAAMLVRIITLLQYSGIFRFLFAITFLNNNTTTPCLPLGCTITA 196
QY 198 SGLDVLPIVSIAGLITGPNRSKAVGPTGVIMNAKRAFAAGVCGPELIDPKGLALVNGT 257
DB 198 SGLDVLPIVSIAGLITGPNRSKAVGPTGVIMNAKRAFAAGVCGPELIDPKGLALVNGT 257
QY 197 SGLDVLPIVSIAGLITGPNRSKAVGPTGVIMNAKRAFAAGVCGPELIDPKGLALVNGT 256
DB 197 SGLDVLPIVSIAGLITGPNRSKAVGPTGVIMNAKRAFAAGVCGPELIDPKGLALVNGT 256
QY 258 AVSGMSASMTLFTANVALLSEVLSAIFAEVMSKPEFTDILTKIKHHPQITEAALIME 317
DB 258 AVSGMSASMTLFTANVALLSEVLSAIFAEVMSKPEFTDILTKIKHHPQITEAALIME 316
QY 257 AVSGMSASMTLFTANVALLSEVLSAIFAEVMSKPEFTDILTKIKHHPQITEAALIME 316
DB 257 AVSGMSASMTLFTANVALLSEVLSAIFAEVMSKPEFTDILTKIKHHPQITEAALIME 316
QY 318 YLIDGSIYVAAKQVHMDIQPKODRYALRTSPQMLGPQILFVIRSSYKMLEREINSNV 377
DB 318 YLIDGSIYVAAKQVHMDIQPKODRYALRTSPQMLGPQILFVIRSSYKMLEREINSNV 377
QY 317 HVLDRSSYVAAKQVHMDIQPKODRYALRTSPQMLGPQILFVIRSSYKMLEREINSNV 376
DB 317 HVLDRSSYVAAKQVHMDIQPKODRYALRTSPQMLGPQILFVIRSSYKMLEREINSNV 376

```

```

QY 438 UNPLIDVSRKALHGNFGSTPLGYSMDNTKLALAIAGLMAHAFSELYNDFTYNNGLPSN 437
DB 438 UNPLIDVSRKALHGNFGSTPLGYSMDNTKLALAIAGLMAHAFSELYNDFTYNNGLPSN 436
QY 438 LSSAPNLSCLYSGKSGEIMASVCSFLQFLANPVNIWVSAEJHJWVNSVSI3CAPTA 497
DB 438 LSSAPNLSCLYSGKSGEIMASVCSFLQFLANPVNIWVSAEJHJWVNSVSI3CAPTA 496
QY 437 LTGGRNPSLDYFGKGAELIMASVCSFLQFLANPVNIWVSAEJHJWVNSVSI3CAPTA 496
DB 437 LTGGRNPSLDYFGKGAELIMASVCSFLQFLANPVNIWVSAEJHJWVNSVSI3CAPTA 496
QY 498 EAVDILKMSSTYLVALUGSIDRHEENKMKSTYKNTVSQVAKKVLGNGELSPRCE 557
DB 498 EAVDILKMSSTYLVALUGSIDRHEENKMKSTYKNTVSQVAKKVLGNGELSPRCE 556
QY 497 EAVDILKMSSTYLVALUGSIDRHEENKMKSTYKNTVSQVAKKVLGNGELSPRCE 556
DB 497 EAVDILKMSSTYLVALUGSIDRHEENKMKSTYKNTVSQVAKKVLGNGELSPRCE 556
QY 558 EKOLLVNDREYFAYIDDVSCSTYPLMOKLROVLVDHAIANNGETKNTNISFOKIAF 617
DB 558 EKOLLVNDREYFAYIDDVSCSTYPLMOKLROVLVDHAIANNGETKNTNISFOKIAF 616
QY 557 EKOLLVNDREYFAYIDDVSCSTYPLMOKLROVLVDHAIANNGETKNTNISFOKIAF 616
DB 557 EKOLLVNDREYFAYIDDVSCSTYPLMOKLROVLVDHAIANNGETKNTNISFOKIAF 616
QY 618 EELKVLPRFEGVPIAYENDLISITNRIKACKSYPLNRPVFEELSGEFTGEKVTSPG 677
DB 618 EELKVLPRFEGVPIAYENDLISITNRIKACKSYPLNRPVFEELSGEFTGEKVTSPG 676
QY 617 EELKVLPRFEGVPIAYENDLISITNRIKACKSYPLNRPVFEELSGEFTGEKVTSPG 676
DB 617 EELKVLPRFEGVPIAYENDLISITNRIKACKSYPLNRPVFEELSGEFTGEKVTSPG 676
QY 678 EEFDRVFTAMCKGQIIDPLELLELGGMMKEPPLPIC 711
DB 678 EEFDRVFTAMCKGQIIDPLELLELGGMMKEPPLPIC 710
QY 677 EEFDRVFTAMCKGQIIDPLELLELGGMMKEPPLPIC 710
DB 677 EEFDRVFTAMCKGQIIDPLELLELGGMMKEPPLPIC 710

RESULT 8
Q9ATN7 PRELIMINARY: PRT: 710 AA.
AC 09ATN7
DT 01-JUN-2001 (TREMblrel, 17, Created)
DT 01-JUN-2001 (TREMblrel, 17, Last sequence update)
DE 01-JUN-2002 (TREMblrel, 21, Last annotation update)
DE phenylalanine ammonia-lyase.
CN PAL1.
OS Agastache rugosa.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots;
OC Asteridae; easterids I; Lamiales; Lamiales; Agastache.
OX NCBI_TaxID=39271;
RN 111
RP SEQUENCE FROM N.A.
RA Kim J.-B., Kim J.-B., Cho K.-J., Moon U., Ha S.-H.;
RT "Molecular cloning and characterization of Phenylalanine Ammonia-
RT lyase(PAL) Related to the Biosynthesis of Rosmarinic acid in Agastache
RT rugosa".
DR Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF326116; AAK15640.1;
DR HSSP: P21310, 1B8F.
DR InterPro: IPR001106; Phc/Hls_NH3lyase.
DR Pfam: PF00221; PAL; 1.
DR TIGRfams: TIGR01226; Phc-am_lyase; 1.
DR PROSITE: PS00488; PAL_HISTIDASE; 1.
KW Lyase.
SQ SEQUENCE 710 AA: 77487 MW; 9403F33B8CFE2958 CRC64.

Query Match 84.8%; Score 3090; DH 10; Length 710;
Best Local Similarity 83.2%; Pred. No. 1.4e-187;
Matches 595; Conservative 54; Mismatches 56; Indels 8; Gaps 4;

QY 3 NGNH---VNGVNVELTK--TPILNVVAEALTSNHLNEVRKVAEPKPVVYLGSETL 56
DB 3 NGNH---VNGVNVELTK--TPILNVVAEALTSNHLNEVRKVAEPKPVVYLGSETL 56
QY 4 NGHGNSGNHNGKANGCHGVKNQDPLNMAAAAPSLKGSNLFVYKVPVVKLGSETL 63
DB 4 NGHGNSGNHNGKANGCHGVKNQDPLNMAAAAPSLKGSNLFVYKVPVVKLGSETL 63
QY 57 TVSGVAGIAAANDSDTVKYLSEARARAVKRSKSSQWVMEKMKGTSTSTGVTTFGATSHPT 116
DB 57 TVSGVAGIAAANDSDTVKYLSEARARAVKRSKSSQWVMEKMKGTSTSTGVTTFGATSHPT 121
QY 64 TTSVVAIVAKKQNA--VAVFLASAPAVKASSQWVMSKPEFTDILTKIKHHPQITEAALIME 121
DB 64 TTSVVAIVAKKQNA--VAVFLASAPAVKASSQWVMSKPEFTDILTKIKHHPQITEAALIME 121
QY 117 TKOGALQKELIFLNAGIFGNGCTETSHUPHSAFRAAMIVRIINTLLQYSGIFRFLFAIT 176
DB 117 TKOGALQKELIFLNAGIFGNGCTETSHUPHSAFRAAMIVRIINTLLQYSGIFRFLFAIT 181
QY 122 TKOGALQKELIFLNAGIFGNGCTETSHUPHSAFRAAMIVRIINTLLQYSGIFRFLFAIT 181
DB 122 TKOGALQKELIFLNAGIFGNGCTETSHUPHSAFRAAMIVRIINTLLQYSGIFRFLFAIT 181
QY 177 TTRFLNNNTTPFLKFTTASGLVPLSYIANLLIGPNRSKAVNITGEVIMNAKRAFAAG 236
DB 177 TTRFLNNNTTPFLKFTTASGLVPLSYIANLLIGPNRSKAVNITGEVIMNAKRAFAAG 236

```


Query Match 83.9%; Score 3056.5; DB 10; Length 718;
 Best local similarity 83.0%; Pred. No. 1,96-185;
 Matches 596, Conservative 47, Mismatches 60, Indels 15, Gaps 4;

QY 2 EN:NNVNGVNNELTK-----NPLNMGVAAELTSHIDEVKKMAVEPRKPVVKLG 53
 8 KNG-HQNGHGLGCKNNNNVSSGDAINMGVMAETLKSHLEEVKKMAVEPRKPVVKLG 66
 DB

QY 54 ELLTVASVNAIAAANDSDTVKVELSEAPPAVKAASDWMWESNNKGTDSYVTTGATG 113
 67 ELLTVASVNAIAAANDSDTVKVELSEAPPAVKAASDWMWESNNKGTDSYVTTGATG 122
 DB

QY 114 HRTKQGAIAQKELIRHINAGIHCNCTETSHLPHSATPAAMIVRINTLLGYSIGREI 173
 123 HRTKNGGALQKELIKRLNNGIFGNGTKSSHLPHSATPAAMIVRINTLLGYSIGREI 182
 DB

QY 174 LEATKELNNITPCLEPRTTASGDIPLASTIAGLTGPPNSKAVPTGTVLNAEKAF 233
 183 LKATKELNNITPCLEPRTTASGDIPLASTIAGLTGPPNSKATGPHGOIIPORAS 242
 DB

QY 234 AAAGVEGGELEDPKESIALVNIITAVSGKASVLEFIAVALLSEVLSAIFAEVMAQKP 293
 243 KPAKF--GFPELOPKESIALVNIITAVSGKASVLEFIAVALLSEVLSAIFAEVMAQKP 300
 DB

QY 294 EFTDHLTHKIKHHPQIIEAALMEVTLDSGVYKAQKVEHMDLPKPKODRYALRTSPQ 353
 301 EFTDHLTHKIKHHPQIIEAALMEHILDSGVYVNAKKEHMDLPKPKODRYALRTSPQ 360
 DB

QY 354 WLSPOLEIVTSSTKMEREINSVNDPLIDVSNKALHGGNFQGTPIGVSMTNRLATAA 413
 361 WLSPOLEIVTSSTKMEREINSVNDPLIDVSNKALHGGNFQGTPIGVSMTNRLATAA 420
 DB

QY 414 IGLMFAQFSELVNDYFNNGPLPSNLSGGNPLDYGFKGSEIIMASVSELOFLANPTN 473
 421 IGLMFAQFSELVNDYFNNGPLPSNLSGGNPLDYGFKGSEIIMASVSELOFLANPTN 480
 DB

QY 474 HVQASQHNODVNSGLISARKTAFAVDILKIMSSYIYALQSIDLRHLENNKSTYKN 533
 481 HVQASQHNODVNSGLISARKTAFAVDILKIMSSYIYALQSIDLRHLENNKSTYKN 540
 DB

QY 534 TVSOVAKKVTIMGVNELHPSRFECEKDLLPVYFPEVYFTEGVSITPCLMKIKUYV 593
 541 TVSOVAKKVTIMGVNELHPSRFECEKDLLPVYFPEVYFTEGVSITPCLMKIKUYV 600
 DB

QY 594 DIALNNGETEKNTSIFOKIATFEELVLLPEKEVEGRIAYENDTSLIPRIKACS 653
 601 DIALNNGETEKNTSIFOKIATFEELVLLPEKEVEGRIAYENDTSLIPRIKACS 660
 DB

QY 654 PLYRIVREHIGROHITGCKYTSIGCEPFDVFTAMCKGQIIDPLKLGCMNGHPPLIC 711
 661 PLYRIVREHIGROHITGCKYTSIGCEPFDVFTAMCKGQIIDPLKLGCMNGHPPLIC 718
 DB

RESULT 11
 024266 PRELIMINARY; PRT, 715 AA.

AC 024266;
 DT 01-JAN-1998 (TREMblrel, 05, created)
 DT 01-JUN-1998 (TREMblrel, 05, last sequence update)
 DT 01-JUN-2002 (TREMblrel, 21, last annotation update)
 DE phenylalanine ammonia-lyase (EC 4.3.1.5).
 OS Populus kitakamensis (Aspen).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Eudicotyledons; Core eudicots; Rosidae;
 OC Eurosid 1; Malpighiales; Salicaceae; Populus.
 ON NCHI_TaxID=1698;
 OX 111
 RN SOURCE: FROM N.A.
 RP OSakabe Y., Ohnubo Y., Kawai S., Katayama Y., Motohashi N.;
 RT "Structures and tissue-specific expression of genes for phenylalanine
 RT ammonia-lyase from a hybrid aspen, Populus kitakamensis.";
 RL Plant Sci. 105:217-226(1995).
 DB FMBI; D3056; HAA21643.1;

DR HSSP: P21310; 188F.
 DR InterPro: IPK001106; Phc/His_NH3lyase.
 DR Pfam: PF00221; PAL; 1.
 DR TrpPfam: TrpP01226; Phc_am_lyase; 1.
 DR PROSITE: PS00488; PAL_HISTIDASE; 1.
 KW lyase.
 SU SOURCE: 715 AA; 77626 MW; 040F098439H740F CMC64;

Query Match 83.7%; Score 3049; DB 10; Length 715;
 Best local similarity 82.7%; Pred. No. 5,66-185;
 Matches 597, Conservative 58, Mismatches 61, Indels 4, Gaps 4;

QY 2 ENGNHGVNNEICI-KDPLNMGVAAELTSHIDEVKKMAVEPRKPVVKLGCTITVSO 60
 6 KNG-YQNGSSSELCIORDPLSKVAAEAKKSHIDEVKKMAVEPRKPVVKLGCTITVSO 64
 DB

QY 61 VASIAAANDSDTVKVELSEAPPAVKAASDWMWESNNKGTDSYVTTGATGSHRTKQ 120
 65 VASIAAANDSDTVKVELSEAPPAVKAASDWMWESNNKGTDSYVTTGATGSHRTKQ 123
 DB

QY 121 GALKKELIRLNGIRGNGTETSHLPHSATPAAMIVRINTLLGYSIGREIETATK 180
 124 GALKKELIRLNGIRGNGTETSHLPHSATPAAMIVRINTLLGYSIGREIETATK 183
 DB

QY 181 LNNITPCLEPRTTASGDIPLASTIAGLTGPPNSKAVPTGTVLNAEKAF 240
 184 LNNITPCLEPRTTASGDIPLASTIAGLTGPPNSKAVPTGTVLNAEKAF 243
 DB

QY 241 GFPELOPKESIALVNIITAVSGKASVLEFIAVALLSEVLSAIFAEVMAQKP 300
 244 GFPELOPKESIALVNIITAVSGKASVLEFIAVALLSEVLSAIFAEVMAQKP 303
 DB

QY 301 HKIKHHPQIIEAALMEVTLDSGVYKAQKVEHMDLPKPKODRYALRTSPQ 360
 304 HKIKHHPQIIEAALMEHILDSGVYVNAKKEHMDLPKPKODRYALRTSPQ 363
 DB

QY 361 VIKSTKMIEREINSVNDPLIDVSNKALHGGNFQGTPIGVSMTNRLATAA 420
 364 VIKSTKMIEREINSVNDPLIDVSNKALHGGNFQGTPIGVSMTNRLATAA 423
 DB

QY 421 QFSEIVNIFNNLPSNLS:RPFSLTYGFKSEIIMASVSELOFLANPTN 480
 424 QFSEIVNIFNNLPSNLS:RPFSLTYGFKSEIIMASVSELOFLANPTN 483
 DB

QY 481 HNDVNSLGLTSARKT-AEAVNDILKIMSSYIYALQSIDLRHLENNKSTYKN 539
 484 HNDVNSLGLTSARKT-AEAVNDILKIMSSYIYALQSIDLRHLENNKSTYKN 543
 DB

QY 540 KKVITMGVNGELHPSRFECEKDLLPVYFPEVYFTEGVSITPCLMKIKUYV 599
 544 KKVITMGVNGELHPSRFECEKDLLPVYFPEVYFTEGVSITPCLMKIKUYV 603
 DB

QY 600 GEFPEKNTSIFOKIATFEELVLLPEKEVEGRIAYENDTSLIPRIKACS 659
 604 GEFPEKNTSIFOKIATFEELVLLPEKEVEGRIAYENDTSLIPRIKACS 663
 DB

QY 660 KEELSRGFEELIGEKVTSIGCEPFDVFTAMCKGQIIDPLKLGCMNGHPPLIC 711
 664 KEELSRGFEELIGEKVTSIGCEPFDVFTAMCKGQIIDPLKLGCMNGHPPLIC 715
 DB

RESULT 12
 094941 PRELIMINARY; PRT, 725 AA.

AC 094941;
 DT 01-DEC-2001 (TREMblrel, 19, created)
 DT 01-DEC-2001 (TREMblrel, 19, last sequence update)
 DT 01-JUN-2002 (TREMblrel, 21, last annotation update)
 DE Ectopic phenylalanine ammonia-lyase.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Eudicotyledons; Core eudicots; Rosidae;
 OC Eurosid 1; Brassicales; Brassicaceae; Arabidopsis.

```

OX NCB1_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Pham P.K., Banh J., Dale J.M., Goldsmith A.D.,
RA Jiang P.X., Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M.,
RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kamita A., Karlin-Neumann G.,
RA Kawai J., Kim C., Koeseha B., Lam B., Lin J., Meyers M.C., Miranda M.,
RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Saitou M., Seki M.,
RA Shinn P., Southwick A., Tracy S.E., Shinozaki K., Davis R.W.,
RA Ecker J.P., Theologis A.,
PT Full length cDNA of gene T1J8.22/Atg37040 (c1:6598547).
RI Submitted (Jul-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY045919; AAK76593.1;
DR InterPro: IPR001106; Phe/His_NH3lyase.
DR Pfam: PF00221; PAL: 1.
DR PRINTS: PR01568; LYMPHOTACTNP
DR TIGRPFAMS: TIGR01226; Phe_am_lyase; 1.
DR PROSITE: PS00488; PAL_HISTIDASE; UNKNOWN_1.
KW lyase.
SQ SEQUENCE 725 AA; 78725 MW; 02626B3BDEF9CE CRC64,
Query Match 83.6%; Score 3045.5; DB 10; Length 725;
Best Local Similarity 81.4%; Pred. No. 9.6e-185;
Matches 569; Conservative 52; Mismatches 66; Indels 17; Gaps 3;
3 NGNHVN---GVNNELC-----IKDPLNMGVAAEALTGSHLDEVKKVAEERKRP 47
4 NGAHKSNGGVDAMLCGGDIKKNVINAEDPLNMGAAAEQMGSHLDEVKRVAAEERKRP 63
48 VYKLGFTLTIVSOVAGIAAANDSTVYKVELSEARACVAKSSIDVWESMKNKCTDSCVTT 107
64 VYNLGSETLTIGVAAISTIGNS--VKVELSETARAGVNASDWMESMKNKGTDSYCVTT 121
108 GFGATSHRRTKGSGALOKELIFELNAGIFGNTEGTSHTLPISATRAAMIVKINTLLAGYS 167
122 GFGATSHRRTKGSGALOKELIFELNAGIFGNTEGTSHTLPISATRAAMIVKINTLLAGYS 181
168 GTFEELLFATIFLNNNTPLPLRGITTAGSDLVPLSYIAGLLTGPENSKAVGPTGEVL 227
182 GTFEELLFATIFLNNNTPLPLRGITTAGSDLVPLSYIAGLLTGPENSKAVGPTGEVL 241
228 NAKKATAAGVCGFEELPKKGLAVNGTAAGSGMASMVLFDANVLAALLSEVLSAIPAE 287
242 TAEFAKKLGISSGFEDLPKGLAVNGTAGSGMASMVLLETNVLSVALEILSAVPAE 301
288 VMGKPEFTDHLTHLKHHPGQIEAAAIEMEYLLDGSDDYVKAOKVHEMDLPQKPKODRYA 347
302 VMGKPEFTDHLTHLKHHPGQIEAAAIEMEYLLDGSDDYVKAOKVHEMDLPQKPKODRYA 361
348 LRTSPWMLGPOUEVIRRSSTKMLEKELINSYNDNPLIDVSRKNAKLGNGFGTPIGVSMDNT 407
362 LRTSPWMLGPOUEVIRRSSTKMLEKELINSYNDNPLIDVSRKNAKLGNGFGTPIGVSMDNT 421
408 RLALIAIGKLMFAOFSELVNDYFNNGLPSNLGGKRNPLSDYGFKGGEIAMAASYCEIQL 467
422 RLALIAIGKLMFAOFSELVNDYFNNGLPSNLGGKRNPLSDYGFKGGEIAMAASYCEIQL 481
468 ANEVTNHVOSADQNDVNSLGLISARKTAEAVDILKMSSTYVALCQSIDLRHLEENM 527
482 ANEVTNHVOSADQNDVNSLGLISARKTAEAVDILKMSSTYVALCQSIDLRHLEENM 541
528 KSTVKNTVQVAKKYLITKYNVNEHLPSRCECDLLRVYDREYVAYIDVCSGTPYLMOK 587
542 RQVKNTVQVAKKYLITKYNVNEHLPSRCECDLLRVYDREYVAYIDVCSGTPYLMOK 601
588 LKQVLDHALINNGEKTNTNTSIFOKATFEELKVLIPKVEGKRLAYENDTISTPNRI 647
602 LKQVLDHALINNGEKTNTNTSIFOKATFEELKVLIPKVEGKRLAYENDTISTPNRI 661
648 KACRSYPLVRYVEBELTGEKVTSPGEEFQVFTAMCKGQJIDPLECTLGGMNGEP 707
662 KACRSYPLVRYVEBELTGEKVTSPGEEFQVFTAMCKGQJIDPLECTLGGMNGEP 721

```

```

OY 708 LPTC 711
DB 722 IFTC 725
RESULT 13
OBRWP4 PRELIMINARY: PRT: 717 AA.
AC OBRWP4: 01-JUN-2002 (TRENBLREL. 21, Created)
DT 01-JUN-2002 (TRENBLREL. 21, Last sequence update)
DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
DE Phenylalanine ammonia-lyase.
GN AT3G53260.
OS Arabidopsis thaliana (mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosidae;
OC eucosids 11; Brassicales; Brassicaceae; Arabidopsids.
OX NCB1_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamita A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Saitou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J.P., Theologis A., Davis R.W.,
RI Submitted (Mar-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY092957; AAK12956.1;
KW lyase.
SQ SEQUENCE 717 AA; 77801 MW; C4C9CBDCDB91A838 CRC64;
Query Match 82.6%; Score 3009; DB 10; Length 717;
Best Local Similarity 81.6%; Pred. No. 2e-182;
Matches 578; Conservative 52; Mismatches 76; Indels 2; Gaps 1;
4 GNVHGVNNELCIKDPLNMGVAAEALTGSHLDEVKKVAEERKRPVYKLGSETLTIVSOVAG 63
12 GRTKVAATVTKTLADPLNMGLAADMKGSHLDEVKKVAEERKRPVYKLGSETLTIGVAA 71
64 IAAANDSTVYKVELSEARACVAKSSIDVWESMKNKCTDSCVTTTFCATSHRRTKKOGAL 123
72 IFTVCS--VKVELSETARAGVNASDWMESMKNKGTDSYCVTTTFCATSHRRTKKOGAL 129
124 OKELIFELNAGIFGNTEGTSHTLPISATRAAMIVKINTLLAGYSIGTFEELFATIFLNN 183
130 OKELIFELNAGIFGNTEGTSHTLPISATRAAMIVKINTLLAGYSIGTFEELFATIFLNN 189
184 NITPCLPIKGITTAGSDLVPLSYIAGLLTGPENSKAVGPTGEVLNAKATAAGVCGGTF 243
190 NITPCLPIKGITTAGSDLVPLSYIAGLLTGPENSKAVGPTGEVLNAKATAAGVCGGTF 249
244 ELQPKKGLAVNGTAAGSGMASMVLFDANVLAALLSEVLSAIPAEVMOGKPEFTDHLTHL 303
250 ELQPKKGLAVNGTAAGSGMASMVLFDANVLAALLSEVLSAIPAEVMOGKPEFTDHLTHL 309
304 KHHFGQIEAAAIEMEYLLDGSDDYVKAOKVHEMDLPQKPKODRYALRTSPWMLGPOUEVIR 363
310 KHHFGQIEAAAIEMEYLLDGSDDYVKAOKVHEMDLPQKPKODRYALRTSPWMLGPOUEVIR 369
364 SSTKMLEKELINSYNDNPLIDVSRKNAKLGNGFGTPIGVSMDNTRLAIAIGKLMFAOFS 423
370 QATKSTIERELNSYNDNPLIDVSRKNAKLGNGFGTPIGVSMDNTRLAIAIAIGKLMFAOFS 429
424 ELVNPYNNGLPSNLGGKRNPLSDYGFKGGEIAMAASYCEIQLANVTNHVOSADQND 483
430 ELVNPYNNGLPSNLGGKRNPLSDYGFKGGEIAMAASYCEIQLANVTNHVOSADQND 489
484 DVNSLGLISARKTAEAVDILKMSSTYVALCQSIDLRHLEENKSTVKNTVQVAKKYL 543
490 DVNSLGLISARKTAEAVDILKMSSTYVALCQSIDLRHLEENKSTVKNTVQVAKKYL 549

```

QY 544 TGCNCGELHSHRCEKDKLHVVDREYFAYIDVCGSTGYVLMQKLVVDAHANNKCFE 603
 DB 550 TGGNSELHSHRCEKDKLHVVDREYFAYIDVCGSTGYVLMQKLVVDAHANNKCFE 603
 QY 604 KNTNTSFQKATFEEELKVLTPKEVEVAVIAYENDLSTPNRKAQSPYLYFVEPEEL 663
 DB 610 KNAVTSIFQKATFEEELKVLTPKEVEVAVIAYENDLSTPNRKAQSPYLYFVEPEEL 663
 QY 664 GRRFLTEKVTSPFEEDKVFYAMKGGVITTELEFVQWNNCEPLPFC 711
 DB 670 GTKLTEKVTSPFEEDKVFYAMKGGVITTELEFVQWNNCEPLPFC 711

RESULT 14

Q94F90 PRELIMINARY: PRT: 687 AA.
 AC 094F90;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Phenylalanine ammonia-lyase 1 (EC 4.3.1.5) (Fragment).
 GN PAL1.
 OS Manihot esculenta (Cassava) (Manihot).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids 1; Malpighiales; Euphorbiaceae; Manihot.
 OX NCBI_TaxID=1983;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Li H., Beecching J.R.;
 RT "Phenylalanine ammonia-lyase 1 gene, PAL1, from cassava";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF383151; AAK6074.1;
 DR InterPro: IPR001106; Phe/His_NH3lyase.
 DR Pfam: PF00221; PAL: 1.
 DR TIGRfams: TIGR01226; phe_am_lyase; 1.
 DR PROSITE: PS00488; PAL_HISTIDASE; UNKNOWN_1.
 KW lyase.
 FT NON_TER 687 687
 SO SEQUENCE 687 AA; 75066 MW; 950A63A1A4BF076 CPe64;

Query Match 82.28; Score 2995; DB 10; Length 687,
 Best Local Similarity 85.88; Pred. No. 1 46-187;
 Matches 576; Conservative 49; Mismatches 44; Indels 2; Gaps 1,

QY 18 DPLNMVAAALTSIHDEYKKAFAEERKPVYKLGSETITVSQVAGIAAANDSTYKVEL 77
 DB 19 DPLNMVAAALTSIHDEYKKAFAEERKPVYKLGSETITVSQVAGIAAANDSTYKVEL 77
 QY 78 SFARAGCAASSIMWMSNMKGCTDYSVGTTCGATSHRRTKOGALOKHILPFLNAGTFC 137
 DB 77 SEEARAVCAASSIMWMSNMKGCTDYSVGTTCGATSHRRTKOGALOKHILPFLNAGTFC 137
 QY 138 NGCTSHHPLHSAIRAMIVRINTILQGYSGIRFELLHAATKRIANNIPPCILPCTTTA 197
 DB 137 NGCTSHHPLHSAIRAMIVRINTILQGYSGIRFELLHAATKRIANNIPPCILPCTTTA 197
 QY 198 SGDLVPLSYTAGLITGRPNKRAVPTSEVLNAKKAFAAVESFPPELAPKRETLVNST 257
 DB 197 SGDLVPLSYTAGLITGRPNKRAVPTSEVLNAKKAFAAVESFPPELAPKRETLVNST 257
 QY 258 AVCGMASMVLTPANVIALSEVLSAIFAVVMOGKTPFTDHLTHKIKHHPGQIEAALIME 317
 DB 257 AVCGMASMVLTPANVIALSEVLSAIFAVVMOGKTPFTDHLTHKIKHHPGQIEAALIME 317
 QY 318 YLLGSGDYVAAAKVHMDPIQKIKODRYALRTSPQWLQPOLIVHSSTKMIEREINSYN 377
 DB 317 YLLGSGDYVAAAKVHMDPIQKIKODRYALRTSPQWLQPOLIVHSSTKMIEREINSYN 377
 QY 378 DNPLIDVSRKALHGNFQSTPTIGVSMNTRTIAIAIGKLMFAQSFELVNDVYNNGLPSN 437
 DB 377 DNPLIDVSRKALHGNFQSTPTIGVSMNTRTIAIAIGKLMFAQSFELVNDVYNNGLPSN 437

QY 438 TSCGRNPSTLYGCKGCHJAMASYCSFLOFAPVNTNHQSAQHONQVNSGLISAKRTA 497
 DB 437 TSCGRNPSTLYGCKGCHJAMASYCSFLOFAPVNTNHQSAQHONQVNSGLISAKRTA 496
 QY 498 EAVDITKLSSTYVALVASTDEHLEFNMKSTVNIVAGVAKVLMVNVNLEHSHRCEK 557
 DB 497 EAVDITKLSSTYVALVASTDEHLEFNMKSTVNIVAGVAKVLMVNVNLEHSHRCEK 556
 QY 558 EKDLKVVDEYVAVYVAVIOWVSTVPLMKCTPQVVTQAINNGFTFKNTNTSFQKATF 617
 DB 557 EKDLKVVDEYVAVYVAVIOWVSTVPLMKCTPQVVTQAINNGFTFKNTNTSFQKATF 616
 QY 618 EEBLKVLTPKEVEVAVIAYENDLSTPNRKAQSPYLYFVEPEELAPKRETLVNST 677
 DB 617 EEBLKVLTPKEVEVAVIAYENDLSTPNRKAQSPYLYFVEPEELAPKRETLVNST 676
 QY 678 EEDRVFTAMC 688
 DB 677 EEDRVFTAMC 687

RESULT 15

Q9XGR3 PRELIMINARY: PRT: 655 AA.
 AC 09XGR3;
 DT 01-NOV-1999 (TREMblrel. 12, Created)
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Phenylalanine ammonia-lyase (Compe).
 OS Vigna unguiculata (Cowpea).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids 1; Fabales; Fabaceae; Papilionoideae; Phaseolaceae; Vigna.
 OX NCBI_TaxID=3917;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kang K.J., Kim S.H., Park Y.I.;
 RT "Isolation of phenylalanine ammonia-lyase gene from Vigna unguiculata,
 RT challenged by cucumber mosaic virus";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF165998; M445384.1;
 DR HSSP: P21310; 188F.
 DR InterPro: IPR001106; Phe/His_NH3lyase.
 DR Pfam: PF00221; PAL, 1.
 DR TIGRfams: TIGR01226; phe_am_lyase; 1.
 DR PROSITE: PS00488; PAL_HISTIDASE; 1.
 KW lyase.
 FT NON_TER 1 1
 SO SEQUENCE 655 AA; 71874 MW; 1F1DC76B49A518BF CPe64;

Query Match 77.4%; Score 4840.5; DB 10; Length 655;
 Best Local Similarity 82.08; Pred. No. 1 5e-170;
 Matches 538; Conservative 57; Mismatches 60; Indels 1; Gaps 1;

QY 56 LTVSQAAGIAANUSQIVKVELSFAARAGCAASSIMWMSNMKGCTDYSVGTTCGATSHR 115
 DB 1 LTVSQAAGIAANUSQIVKVELSFAARAGCAASSIMWMSNMKGCTDYSVGTTCGATSHR 115
 QY 116 KIKQVSLQKELLPFLNAGTFSNRTSHHPLHSAIRAMIVRINTILQGYSGIRFELLHA 175
 DB 115 KIKQVSLQKELLPFLNAGTFSNRTSHHPLHSAIRAMIVRINTILQGYSGIRFELLHA 174
 QY 176 AITKELNNNTTPLLKGTITASGDLVPLSYTAGLITGRPNKRAVPTSEVLNAKKAFAA 235
 DB 175 AITKELNNNTTPLLKGTITASGDLVPLSYTAGLITGRPNKRAVPTSEVLNAKKAFAA 234
 QY 236 AGVAGCGFELQPKRGALVAVGAVCGMASMVLTPANVIALSEVLSAIFAVVMOGKTPF 295
 DB 235 AGVAGCGFELQPKRGALVAVGAVCGMASMVLTPANVIALSEVLSAIFAVVMOGKTPF 294
 QY 296 TDHILTHKIKHHPGQIEAALIMEYITLDGSDVKAQVHMDPIQKIKODRYALRTSPQWL 355
 DB 295 TDHILTHKIKHHPGQIEAALIMEYITLDGSDVKAQVHMDPIQKIKODRYALRTSPQWL 354
 QY 355 TDHILTHKIKHHPGQIEAALIMEYITLDGSDVKAQVHMDPIQKIKODRYALRTSPQWL 354
 DB 354 TDHILTHKIKHHPGQIEAALIMEYITLDGSDVKAQVHMDPIQKIKODRYALRTSPQWL 353

```

QY 356 GFOIEYIRSTKMEIEEINSVNDNPLIDVSRKALHGGNFUGFPIGVSMNTRLAIAIG 415
Db 300 GPQIEYIRHATKMIEREINSVNDNPLIDVSRKALHGGNFUGFPIGVSMNTRLAIASIG 359
QY 416 KLMFAQFSELVNFYNNMSPNLG3RNP5LDYGFK3GEIAMASYCSELOFLANPVTNHV 475
Db 360 KLMFAQFSELVNDFYNNMSPNLG3RNP5LDYGFK3GEIAMASYCSELOFLANPVTNHV 419
QY 476 QSAFOHNOUVNSGLISAFTAAVDILKMSSTYVALCOSIDLRIEENKSTVKNV 535
Db 420 QSAFOHNOUVNSGLVSSSKTAEEELKMSSTFLVALCOAIDRLHEENMKNAVNAV 479
QY 536 SOYAKKVLTMGVNGELHPSRCECKDLRVVDREYFAYIDVCSGTYPLMOKLRQVLVDH 595
Db 480 SOYAKKVLTTGNGELHPSRCECKDLKIVDREYFAYIDPCSVTYPLMOKLRQVLVDH 539
QY 596 ALNNGETEKNTSIPUKIATFEELKVLLEKEVGRIAYENDTSLIPNRIKACRSYPL 655
Db 540 ALNNGEKAASSSTSIFFKIRAFEELITLLPREVENARVEVENGNSSIPNRIKACRSYPL 599
QY 656 YRFVREELGRGLTGKVTSPGEFDRVFTAMCKGQIIDPLLEGLGWNCEPLPIC 711
Db 600 YRFVREELSTSLQYGEKVASPGECDKVFTALCEGKFTIDPMDCILKKMNGSPLPIC 655

```

Search completed: March 29, 2003, 02:12:14
 Job time : 67 secs